SEQUENCE LISTING

<110> Wesley, Susan V.
Waterhouse, Peter
Helliwell,Christopher A.

<120> Method and means for producing efficient silencing constructs using recombinational cloning

<130> HELLGA

<160> 26

<170> PatentIn version 3.1

<210> 1
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attB1

<400> 1
agcctgcttt tttgtacaaa cttgt 25

<210> 2
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attB2

<400> 2
agcctgcttt cttgtacaaa cttgt 25

<210> 3
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attB3

<400> 3
acccagcttt cttgtacaaa cttgt 25

<210> 4
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attR1

<400> 4
gttcagcttt tttgtacaaa cttgt 25

<210> 5
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attR2

<400> 5
gttcagcttt cttgtacaaa cttgt 25

<210> 6
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attR3

<400> 6
gttcagcttt cttgtacaaa gttgg 25

<210> 7
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attL1

<400> 7
agcctgcttt tttgtacaaa gttgg 25

<210> 8
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attL2

<400> 8
agcctgcttt cttgtacaaa gttgg 25

<210> 9
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attL3

<400> 9
acccagcttt cttgtacaaa gttgg 25

<210> 10
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attP1

<400> 10
gttcagcttt tttgtacaaa gttgg 25

<210> 11
<211> 25
<212> DNA
<213> Artificial sequence

<220>
<223> core sequence of recombination site attP2,P3

<400> 11
gttcagcttt cttgtacaaa gttgg 25

<210> 12
<211> 1188
<212> DNA
<213> Artificial sequence

<220>
<223> cDNA sequence of the Arabidopsis thaliana chalcone synthase codin
g region

<400> 12
atggtgatgg ctggtgcttc ttctttggat gagatcagac aggctcagag agctgatgga 60
cctgcaggca tcttggctat tggcactgct aaccctgaga accatgtgct tcaggcggag 120
tatcctgact actacttccg catcaccaac agtgaacaca tgaccgacct caaggagaag 180
ttcaagcgca tgtgcgacaa gtcgacaatt cggaaacgtc acatgcatct gacggaggaa 240
ttcctcaagg aaaacccaca catgtgtgct tacatggctc cttctctgga caccagacag 300
gacatcgtgg tggtcgaagt ccctaagcta ggcaaagaag cggcagtgaa ggccatcaag 360
gagtggggcc agcccaagtc aaagatcact catgtcgtct tctgcactac ctccggcgtc 420
gacatgcctg gtgctgacta ccagctcacc aagcttcttg gtctccgtcc ttccgtcaag 480
cgtctcatga tgtaccagca aggttgcttc gccggcggta ctgtcctccg tatcgctaag 540
gatctcgccg agaacaaccg tggagcacgt gtcctcgttg tctgctctga gatcacagcc 600
gttaccttcc gtggtccctc tgacacccac cttgactccc tcgtcggtca ggctcttttc 660
agtgatggcg ccgccgcact cattgtgggg tcggaccctg acacatctgt cggagagaaa 720
cccatctttg agatggtgtc tgccgctcag accatccttc cagactctga tggtgccata 780
gacggacatt tgagggaagt tggtctcacc ttccatctcc tcaaggatgt tcccggcctc 840
atctccaaga acattgtgaa gagtctagac gaagcgttta aacctttggg gataagtgac 900
tggaactccc tcttctggat agcccaccct ggaggtccag cgatcctaga ccaggtggag 960
ataaagctag gactaaagga agagaagatg agggcgacac gtcacgtgtt gagcgagtat 1020
ggaaacatgt cgagcgcgtg cgttctcttc atactagacg agatgaggag gaagtcagct 1080
aaggatggtg tggccacgac aggagaaggg ttggagtggg gtgtcttgtt tggtttcgga 1140
ccaggtctca ctgttgagac agtcgtcttg cacagcgttc ctctctaa 1188

<210> 13
<211> 18691
<212> DNA
<213> Artificial sequence

<220>
<223> acceptor vector pHELLSGATE

<220>
<221> misc_feature
<222> (7922)..(9985)
<223> spectinomycin resistance

<220>
<221> misc_feature
<222> (10706)..(11324)
<223> right T-DNA border fragment

<220>
<221> misc_feature
<222> (11674)..(13019)
<223> CaMV35S promoter fragment

<220>
<221> misc_feature
<222> (17890)..(17659)
<223> attP1 recombination site (complement)

<220>
<221> misc_feature
<222> (17610)..(16855)
<223> ccdB selection marker (complement)

<220>
<221> misc_feature
<222> (16551)..(16319)
<223> attP2 recombination site (complement)

<220>
<221> misc_feature
<222> (14660)..(16258)
<223> pdk2 intron 2

<220>
<221> misc_feature
<222> (15002)..(15661)
<223> chloramphenicol resistance gene

<220>
<221> misc_feature
<222> (14387)..(14619)
<223> attP2 recombination site

<220>
<221> misc_feature
<222> (13675)..(13980)
<223> ccdB selection marker (complement)

<220>
<221> misc_feature
<222> (13048)..(13279)
<223> attP1 recombination site

<220>
<221> misc_feature
<222> (17922)..(18687)
<223> octopine synthase gene terminator region

<220>
<221> misc_feature
<222> (264)..(496)
<223> nopaline synthase gene promoter

<220>
<221> misc_feature
<222> (497)..(1442)
<223> nptII coding region

<220>
<221> misc_feature
<222> (1443)..(2148)
<223> nopaline synthase gene terminator

<220>
<221> misc_feature
<222> (2149)..(2706)
<223> a left T-DNA border region

<400> 13

```
ggccgcacta gtgatatccc gcggccatgg cggccgggag catgcgacgt cgggcccaat     60
tcgccctata gtgagtcgta ttacaattca ctggccgtcg ttttacaacg tcgtgactgg    120
gaaaaccctg gcgttaccca acttaatcgc cttgcagcac atccccctttcgccagctgg    180
cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag cctgaatggc    240
gaatggaaat tgtaaacgtt aatgggtttc tggagtttaa tgagctaagc acatacgtca    300
gaaaccatta ttgcgcgttc aaaagtcgcc taaggtcact atcagctagc aaatatttct    360
tgtcaaaaat gctccactga cgttccataa attcccctcg gtatccaatt agagtctcat    420
attcactctc aatccaaata atctgcaatg gcaattacct tatccgcaac ttctttacct    480
atttccgccc ggatccgggc aggttctccg gccgcttggg tggagaggct attcggctat    540
gactgggcac aacagacaat cggctgctct gatgccgccg tgttccggct gtcagcgcag    600
gggcgcccgg ttcttttttgt caagaccgac ctgtccggtg ccctgaatga actgcaggac    660
gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac    720
gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg gcaggatctc    780
ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc aatgcggcgg    840
ctgcatacgc ttgatccggc tacctgccca ttcgaccacc aagcgaaaca tcgcatcgag    900
cgagcacgta ctcggatgga agccggtctt gtcgatcagg atgatctgga cgaagagcat    960
caggggctcg cgccagccga actgttcgcc aggctcaagg cgcgcatgcc cgacggcgag   1020
gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc   1080
ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg   1140
ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg   1200
ctttacggta tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct tcttgacgag   1260
ttcttctgag cgggactctg gggttcgaaa tgaccgacca agcgacgccc aacctgccat   1320
cacgagattt cgattccacc gccgccttct atgaaaggtt gggcttcgga atcgttttcc   1380
gggacgccgg ctggatgatc ctccagcgcg gggatctcat gctggagttc ttcgcccacc   1440
ccgatccaac acttacgttt gcaacgtcca agagcaaata gaccacgaac gccggaaggt   1500
tgccgcagcg tgtggattgc gtctcaattc tctcttgcag gaatgcaatg atgaatatga   1560
tactgactat gaaactttga gggaatactg cctagcaccg tcacctcata acgtgcatca   1620
tgcatgccct gacaacatgg aacatcgcta tttttctgaa gaattatgct cgttggagga   1680
tgtcgcggca attgcagcta ttgccaacat cgaactaccc ctcacgcatg cattcatcaa   1740
tattattcat gcggggaaag gcaagattaa tccaactggc aaatcatcca gcgtgattgg   1800
taacttcagt tccagcgact tgattcgttt tggtgctacc cacgttttca ataaggacga   1860
gatggtggag taaagaagga gtgcgtcgaa gcagatcgtt caaacatttg gcaataaagt   1920
ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat   1980
tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag atgggttttt   2040
atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca   2100
aactaggata aattatcgcg cgcggtgtca tctatgttac tagatcgaat taattccagg   2160
cggtgaaggg caatcagctg ttgcccgtct cactggtgaa aagaaaaacc accccagtac   2220
attaaaaacg tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa   2280
tatatcctgc caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca   2340
ctcgatacag gcagcccatc agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg   2400
cagactttgc tcatgttacc gatgctattc ggaagaacgg caactaagct gccgggtttg   2460
```

aaacacggat gatctcgcgg agggtagcat gttgattgta acgatgacag agcgttgctg 2520
cctgtgatca aatatcatct ccctcgcaga gatccgaatt atcagccttc ttattcattt 2580
ctcgcttaac cgtgacaggc tgtcgatctt gagaactatg ccgacataat aggaaatcgc 2640
tggataaagc cgctgaggaa gctgagtggc gctatttctt tagaagtgaa cgttgacgat 2700
gtcgacggat cttttccgct gcataaccct gcttcggggt cattatagcg attttttcgg 2760
tatatccatc ctttttcgca cgatatacag gattttgcca aagggttcgt gtagactttc 2820
cttggtgtat ccaacggcgt cagccgggca ggataggtga agtaggccca cccgcgagcg 2880
ggtgttcctt cttcactgtc ccttattcgc acctggcggt gctcaacggg aatcctgctc 2940
tgcgaggctg gccggctacc gccggcgtaa cagatgaggg caagcggatg gctgatgaaa 3000
ccaagccaac caggggtgat gctgccaact tactgattta gtgtatgatg gtgtttttga 3060
ggtgctccag tggcttctgt ttctatcagc tgtccctcct gttcagctac tgacggggtg 3120
gtgcgtaacg gcaaaagcac cgccggacat cagcgctatc tctgctctca ctgccgtaaa 3180
acatggcaac tgcagttcac ttacaccgct tctcaacccg gtacgcacca gaaaatcatt 3240
gatatggcca tgaatggcgt tggatgccgg gcaacagccc gcattatggg cgttggcctc 3300
aacacgattt tacgtcactt aaaaaactca ggccgcagtc ggtaacctcg cgcatacagc 3360
cgggcagtga cgtcatcgtc tgcgcggaaa tggacgaaca gtggggctat gtcggggcta 3420
aatcgcgcca gcgctggctg ttttacgcgt atgacagtct ccggaagacg gttgttgcgc 3480
acgtattcgg tgaacgcact atggcgacgc tggggcgtct tatgagcctg ctgtcaccct 3540
ttgacgtggt gatatggatg acggatggct ggccgctgta tgaatcccgc ctgaagggaa 3600
agctgcacgt aatcagcaag cgatatacgc agcgaattga gcggcataac ctgaatctga 3660
ggcagcacct ggcacggctg ggacggaagt cgctgtcgtt ctcaaaatcg gtggagctgc 3720
atgacaaagt catcgggcat tatctgaaca taaaacacta tcaataagtt ggagtcatta 3780
cccaaccagg aagggcagcc cacctatcaa ggtgtactgc cttccagacg aacgaagagc 3840
gattgaggaa aaggcggcgg cggccggcat gagcctgtcg gcctacctgc tggccgtcgg 3900
ccagggctac aaaatcacgg gcgtcgtgga ctatgagcac gtccgcgagc tggcccgcat 3960
caatggcgac ctgggccgcc tgggcggcct gctgaaactc tggctcaccg acgacccgcg 4020
cacggcgcgg ttcggtgatg ccacgatcct cgccctgctg gcgaagatcg aagagaagca 4080
ggacgagctt ggcaaggtca tgatgggcgt ggtccgcccg agggcagagc catgactttt 4140
ttagccgcta aaacggccgg ggggtgcgcg tgattgccaa gcacgtcccc atgcgctcca 4200
tcaagaagag cgacttcgcg gagctggtat tcgtgcaggg caagattcgg aataccaagt 4260
acgagaagga cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc 4320
tggacaccaa ggcaccaggc gggtcaaatc aggaataagg gcacattgcc ccggcgtgag 4380
tcggggcaat cccgcaagga gggtgaatga atcggacgtt tgaccggaag gcatacaggc 4440
aagaactgat cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg 4500
tcatgcgtgc gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg 4560
ccaagatcga gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg 4620
ccgtggagcg ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag tcgatgacca 4680
tcgacacgcg aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa 4740
aacaggtcag cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg 4800
aaatgcagct ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa 4860
acgacacggc ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc 4920
tgcaaaacaa ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg 4980
agctgcgggc cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca 5040
cccctatcgg cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt 5100
cgatcaatgg ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg 5160
cgatgggctt cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct 5220
tccgcgtcct ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa 5280
tcgtcgtgct gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc 5340
tgtcgccgac ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc 5400
tcaagctgga aaccttccgc ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc 5460
gcgagcaggt cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct 5520
gggtcaatga tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg 5580
ggggttcagc agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca 5640
cttgcttcgc tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag 5700
aggattaaaa ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt 5760
gcaggatttc cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc 5820
cgtttacgag cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc 5880
cgtggcattc ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga 5940
ggacggcccc aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca 6000
gcgaggccga ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt 6060
gatgatcgtc cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc 6120
acttaatatt tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg 6180
ggtcgcggcg acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct 6240

gctaggtagc ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt 6300
ggcgctgttg gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct 6360
ggcgggggcg gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt 6420
gcctctgctc acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt 6480
agctttagtg tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc 6540
gtggctcggc ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg 6600
actcgaacct acagttgttt ccttactggg ctttctcagc cgggatggcg ctaagaagct 6660
attgccgccg atcttcatat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac 6720
cgcatcaggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg 6780
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat 6840
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 6900
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacaa aaatcgacgc 6960
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 7020
agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt 7080
ctcccttcgg gaagcgtggc gctttctcaa tgctcacgct gtaggtatct cagttcggtg 7140
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc 7200
gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg 7260
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc 7320
ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg 7380
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc 7440
gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatat 7500
caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt 7560
taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa 7620
aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa 7680
tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc 7740
tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct 7800
gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca 7860
gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt 7920
aaacaagtgg cagcaacgga ttcgcaaacc tgtcacgcct tttgtgccaa aagccgcgcc 7980
aggtttgcga tccgctgtgc caggcgttag gcgtcatatg aagatttcgg tgatccctga 8040
gcaggtggcg gaaacattgg atgctgagaa ccatttcatt gttcgtgaag tgttcgatgt 8100
gcacctatcc gaccaaggct ttgaactatc taccagaagt gtgagcccct accggaagga 8160
ttacatctcg gatgatgact ctgatgaaga ctctgcttgc tatggcgcat tcatcgacca 8220
agagcttgtc gggaagattg aactcaactc aacatggaac gatctagcct ctatcgaaca 8280
cattgttgtg tcgcacacgc accgaggcaa aggagtcgcg cacagtctca tcgaatttgc 8340
gaaaaagtgg gcactaagca gacagctcct tggcatacga ttagagacac aaacgaacaa 8400
tgtacctgcc tgcaatttgt acgcaaaatg tggctttact ctcggcggca ttgacctgtt 8460
cacgtataaa actagacctc aagtctcgaa cgaaacagcg atgtactggt actggttctc 8520
gggagcacag gatgacgcct aacaattcat tcaagccgac accgcttcgc ggcgcggctt 8580
aattcaggag ttaaacatca tgagggaagc ggtgatcgcc gaagtatcga ctcaactatc 8640
agaggtagtt ggcgtcatcg agcgccatct cgaaccgacg ttgctggccg tacatttgta 8700
cggctccgca gtggatggcg gcctgaagcc acacagtgat attgatttgc tggttacggt 8760
gaccgtaagg cttgatgaaa caacgcggcg agctttgatc aacgaccttt tggaaacttc 8820
ggcttccccc ggagagagcg agattctccg cgctgtagaa gtcaccattg ttgtgcacga 8880
cgacatcatt ccgtggcgtt atccagctaa gcgcgaactg caatttggag aatggcagcg 8940
caatgacatt cttgcaggta tcttcgagcc agccacgatc gacattgatc tggctatctt 9000
gctgacaaaa gcaagagaac atagcgttgc cttggtaggt ccagcggcgg aggaactctt 9060
tgatccggtt cctgaacagg atctatttga ggcgctaaat gaaaccttaa cgctatggaa 9120
ctcgccgccc gactgggctg gcgatgagcg aaatgtagtg cttacgttgt cccgcatttg 9180
gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc gctgccgact gggcaatgga 9240
gcgcctgccg gcccagtatc agcccgtcat acttgaagct aggcaggctt atcttggaca 9300
agaagatcgc ttggcctcgc gcgcagatca gttggaagaa tttgttcact acgtgaaagg 9360
cgagatcacc aaggtagtcg gcaaataatg tctaacaatt cgttcaagcc gacgccgctt 9420
cgcggcgcgg cttaactcaa gcgttagaga gctggggaag actatgcgcg atctgttgaa 9480
ggtggttcta agcctcgtac ttgcgatggc atcggggcag gcacttgctg acctgccaat 9540
tgttttagtg gatgaagctc gtcttcccta tgactactcc ccatccaact acgacatttc 9600
tccaagcaac tacgacaact ccataagcaa ttacgacaat agtccatcaa attacgacaa 9660
ctctgagagc aactacgata atagttcatc caattacgac aatagtcgca acggaaatcg 9720
taggcttata tatagcgcaa atgggtctcg cactttcgcc ggctactacg tcattgccaa 9780
caatgggaca acgaacttct tttccacatc tggcaaaagg atgttctaca ccccaaaagg 9840
ggggcgcggc gtctatggcg gcaaagatgg gagcttctgc ggggcattgg tcgtcataaa 9900
tggccaattt tcgcttgccc tgacagataa cggcctgaag atcatgtatc taagcaacta 9960
gcctgctctc taataaaatg ttaggagctt ggctgccatt tttggggtga ggccgttcgc 10020

```
ggccgagggg cgcagcccct ggggggatgg gaggcccgcg ttagcgggcc gggagggttc  10080
gagaaggggg ggcacccccc ttcggcgtgc gcggtcacgc gccagggcgc agccctggtt  10140
aaaaacaagg tttataaata ttggtttaaa agcaggttaa aagacaggtt agcggtggcc  10200
gaaaaacggg cggaaaccct tgcaaatgct ggatttttctg cctgtggaca gcccctcaaa  10260
tgtcaatagg tgcgcccctc atctgtcagc actctgcccc tcaagtgtca aggatcgcgc  10320
ccctcatctg tcagtagtcg cgcccctcaa gtgtcaatac cgcagggcac ttatccccag  10380
gcttgtccac atcatctgtg ggaaactcgc gtaaaatcag gcgttttcgc cgatttgcga  10440
ggctggccag ctccacgtcg ccggccgaaa tcgagcctgc ccctcatctg tcaacgccgc  10500
gccgggtgag tcggcccctc aagtgtcaac gtccgccccc catctgtcag tgagggccaa  10560
gttttccgcg aggtatccac aacgccggcg gccggccgcg gtgtctcgca cacggcttcg  10620
acggcgtttc tggcgcgttt gcagggccat agacggccgc cagcccagcg gcgagggcaa  10680
ccagcccggt gagcgtcgga aagggtcgac atcttgctgc gttcggatat tttcgtggag  10740
ttcccgccac agacccggat tgaaggcgag atccagcaac tcgcgccaga tcatcctgtg  10800
acggaacttt ggcgcgtgat gactggccag gacgtcggcc gaaagagcga caagcagatc  10860
acgattttcg acagcgtcgg atttgcgatc gaggattttt cggcgctgcg ctacgtccgc  10920
gaccgcgttg agggatcaag ccacagcagc ccactcgacc ttctagccga cccagacgag  10980
ccaagggatc tttttggaat gctgctccgt cgtcaggctt tccgacgttt gggtggttga  11040
acagaagtca ttatcgtacg gaatgccagc actcccgagg ggaaccctgt ggttggcatg  11100
cacatacaaa tggacgaacg gataaacctt ttcacgccct tttaaatatc cgttattcta  11160
ataaacgctc ttttctctta ggtttacccg ccaatatatc ctgtcaaaca ctgatagttt  11220
aaactgaagg cgggaaacga caatctgatc atgagcggag aattaaggga gtcacgttat  11280
gaccccccgcc gatgacgcgg gacaagccgt tttacgtttg gaactgacag aaccgcaacg  11340
attgaaggag ccactcagcc ccaatacgca aaccgcctct ccccgcgcgt tggccgattc  11400
attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa  11460
ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg cttccggctc  11520
gtatgttgtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg  11580
attacgccaa gctatttagg tgacactata gaatactcaa gctatgcatc caacgcgttg  11640
ggagctctcc catatcgacc tgcaggcggc cgctcgacga attaattcca atcccacaaa  11700
aatctgagct taacagcaca gttgctcctc tcagagcaga atcgggtatt caacaccctc  11760
atatcaacta ctacgttgtg tataacggtc cacatgccgg tatatacgat gactggggtt  11820
gtacaaaggc ggcaacaaac ggcgttcccg gagttgcaca caagaaattt gccactatta  11880
cagaggcaag agcagcagct gacgcgtaca caacaagtca gcaaacagac aggttgaact  11940
tcatccccaa aggagaagct caactcaagc ccaagagctt tgctaaggcc ctaacaagcc  12000
caccaaagca aaaagcccac tggctcacgc taggaaccaa aaggcccagc agtgatccag  12060
ccccaaaaga gatctccttt gccccggaga ttacaatgga cgatttcctc tatctttacg  12120
atctaggaag gaagttcgaa ggtgaaggtg acgacactat gttcaccact gataatgaga  12180
aggttagcct cttcaatttc agaaagaatg ctgacccaca gatggttaga gaggcctacg  12240
cagcaggtct catcaagacg atctacccga gtaacaatct ccaggagatc aaataccttc  12300
ccaagaaggt taaagatgca gtcaaaagat tcaggactaa ttgcatcaag aacacagaga  12360
aagacatatt tctcaagatc agaagtacta ttccagtatg gacgattcaa ggcttgcttc  12420
ataaaccaag gcaagtaata gagattggag tctctaaaaa ggtagttcct actgaatcta  12480
aggccatgca tggagtctaa gattcaaatc gaggatctaa cagaactcgc cgtgaagact  12540
ggcgaacagt tcatacagag tcttttacga ctcaatgaca agaagaaaat cttcgtcaac  12600
atggtggagc acgacactct ggtctactcc aaaaatgtca aagatacagt ctcagaagac  12660
caaagggcta ttgagacttt tcaacaaagg ataatttcgg gaaacctcct cggattccat  12720
tgcccagcta tctgtcactt catcgaaagg acagtagaaa aggaaggtgg ctcctacaaa  12780
tgccatcatt gcgataaagg aaaggctatc attcaagatc tctctgccga cagtggtccc  12840
aaagatggac ccccacccac gaggagcatc gtggaaaaag aagacgttcc aaccacgtct  12900
tcaaagcaag tggattgatg tgacatctcc actgacgtaa gggatgacgc acaatcccac  12960
tatccttcgc aagacccttc ctctatataa ggaagttcat ttcatttgga gaggacacgc  13020
tcgaggctag catggatctc gggccccaaa taatgatttt attttgactg atagtgacct  13080
gttcgttgca acaaattgat gagcaatgct tttttataat gccaactttg tacaaaaaag  13140
ctgaacgaga aacgtaaaat gatataaata tcaatatatt aaattagatt ttgcataaaa  13200
aacagactac ataatactgt aaaacacaac atatccagtc actatgaatc aactacttag  13260
atggtattag tgacctgtag tcgaccgaca gccttccaaa tgttcttcgg gtgatgctgc  13320
caacttagtc gaccgacagc cttccaaatg ttcttctcaa acggaatcgt cgtatccagc  13380
ctactcgcta ttgtcctcaa tgccgtatta aatcataaaa agaaataaga aaaagaggtg  13440
cgagcctctt ttttgtgtga caaaataaaa acatctacct attcatatac gctagtgtca  13500
tagtcctgaa aatcatctgc atcaagaaca atttcacaac tcttatactt ttctcttaca  13560
agtcgttcgg cttcatctgg attttcagcc tctatactta ctaaacgtga taaagtttct  13620
gtaatttcta ctgtatcgac ctgcagactg gctgtgtata agggagcctg acatttatat  13680
tccccagaac atcaggttaa tggcgttttt gatgtcattt tcgcggtggc tgagatcagc  13740
cacttcttcc ccgataacgg agaccggcac actggccata tcggtggtca tcatgcgcca  13800
```

gctttcatcc ccgatatgca ccaccgggta aagttcacgg gagactttat ctgacagcag 13860
acgtgcactg gccaggggga tcaccatccg tcgcccgggc gtgtcaataa tatcactctg 13920
tacatccaca aacagacgat aacggctctc tcttttatag gtgtaaacct taaactgcat 13980
ttcaccagtc cctgttctcg tcagcaaaag agccgttcat ttcaataaac cgggcgacct 14040
cagccatccc ttcctgattt tccgctttcc agcgttcggc acgcagacga cgggcttcat 14100
tctgcatggt tgtgcttacc agaccggaga tattgacatc atatatgcct tgagcaactg 14160
atagctgtcg ctgtcaactg tcactgtaat acgctgcttc atagcacacc tctttttgac 14220
atacttcggg tagtgccgat caacgtctca ttttcgccaa aagttggccc agggcttccc 14280
ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg tcacaggtat 14340
ttattcggcg caaagtgcgt cgggtgatgc tgccaactta gtcgactaca ggtcactaat 14400
accatctaag tagttgattc atagtgactg gatatgttgt gttttacagt attatgtagt 14460
ctgtttttta tgcaaaatct aatttaatat attgatattt atatcatttt acgtttctcg 14520
ttcagctttc ttgtacaaag ttggcattat aagaaagcat tgcttatcaa tttgttgcaa 14580
cgaacaggtc actatcagtc aaaataaaat cattatttgc catccagctg cagctcctcg 14640
aggaattcgg taccccaatt ggtaaggaaa taattatttt cttttttcct tttagtataa 14700
aatagttaag tgatgttaat tagtatgatt ataataatat agttgttata attgtgaaaa 14760
aataatttat aaatatattg tttacataaa caacatagta atgtaaaaaa atatgacaag 14820
tgatgtgtaa gacgaagaag ataaaagttg agagtaagta tattattttt aatgaatttg 14880
atcgaacatg taagatgata tacggccggt aagaggttcc aactttcacc ataatgaaat 14940
aagatcacta ccgggcgtat tttttgagtt atcgagattt tcaggagcta aggaagctaa 15000
aatggagaaa aaaatcactg gatataccac cgttgatata tcccaatggc atcgtaaaga 15060
acattttgag gcatttcagt cagttgctca atgtacctat aaccagaccg ttcagctgga 15120
tattacggcc tttttaaaga ccgtaaagaa aaataagcac aagttttatc cggcctttat 15180
tcacattctt gcccgcctga tgaatgctca tccggaattc cgtatggcaa tgaaagacgg 15240
tgagctggtg atatgggata gtgttcaccc ttgttacacc gttttccatg agcaaactga 15300
aacgttttca tcgctctgga gtgaatacca cgacgatttc cggcagtttc tacacatata 15360
ttcgcaagat gtggcgtgtt acggtgaaaa cctggcctat ttccctaaag ggtttattga 15420
gaatatgttt ttcgtctcag ccaatccctg ggtgagtttc accagttttg atttaaacgt 15480
ggccaatatg gacaacttct tcgcccccgt tttcaccatg ggcaaatatt atacgcaagg 15540
cgacaaggtg ctgatgccgc tggcgattca ggttcatcat gccgtctgtg atggcttcca 15600
tgtcggcaga atgcttaatg aattacaaca gtactgcgat gagtggcagg gcggggcgta 15660
atcgcgtgga tccggcttac taaaagccag ataacagtat gcgtatttgc gcgctgattt 15720
ttgcggtata agaatatata ctgatatgtc gggcccataa tagtaattct agctggtttg 15780
atgaattaaa tatcaatgat aaaatactat agtaaaaata agaataaata aattaaaata 15840
atattttttt atgattaata gtttattata taattaaata tctataccat tactaaatat 15900
tttagtttaa aagttaataa atattttgtt agaaattcca atctgcttgt aatttatcaa 15960
taaacaaaat attaaataac aagctaaagt aacaaataat atcaaactaa tagaaacagt 16020
aatctaatgt aacaaaacat aatctaatgc taatataaca aagcgcaaga tctatcattt 16080
tatatagtat tattttcaat caacattctt attaatttct aaataatact tgtagtttta 16140
ttaacttcta aatggattga ctattaatta aatgaattag tcgaacatga ataaacaagg 16200
taacatgata gatcatgtca ttgtgttatc attgatctta catttggatt gattacagtt 16260
gggaaattgg gttcgaaatc gataagcttg gatcctctag agagctgcag ctggatggca 16320
aataatgatt ttattttgac tgatagtgac ctgttcgttg caacaaattg ataagcaatg 16380
ctttcttata atgccaactt tgtacaagaa agctgaacga gaaacgtaaa atgatataaa 16440
tatcaatata ttaaattaga ttttgcataa aaaacagact acataatact gtaaaacaca 16500
acatatccag tcactatgaa tcaactactt agatggtatt agtgacctgt agtcgactaa 16560
gttggcagca tcacccgacg cactttgcgc cgaataaata cctgtgacgg aagatcactt 16620
cgcagaataa ataaatcctg gtgtccctgt tgataccggg aagccctggg ccaacttttg 16680
gcgaaaatga gacgttgatc ggcactaccc atttcacaac tcttatactt ttctcttaca 16740
agtcgttcgg cttcatctgg attttcagcc tctatactta ctaaacgtga taaagtttct 16800
gtaatttcta ctgtatcgac ctgcagactg gctgtgtata agggagcctg acatttatat 16860
tccccagaac atcaggttaa tggcgttttt gatgtcattt tcgcggtggc tgagatcagc 16920
cacttcttcc ccgataacgg agaccggcac actggccata tcggtggtca tcatgcgcca 16980
gctttcatcc ccgatatgca ccaccgggta aagttcacgg gagactttat ctgacagcag 17040
acgtgcactg gccaggggga tcaccatccg tcgcccgggc gtgtcaataa tatcactctg 17100
tacatccaca aacagacgat aacggctctc tcttttatag gtgtaaacct taaactgcat 17160
ttcaccagtc cctgttctcg tcagcaaaag agccgttcat ttcaataaac cgggcgacct 17220
cagccatccc ttcctgattt tccgctttcc agcgttcggc acgcagacga cgggcttcat 17280
tctgcatggt tgtgcttacc agaccggaga tattgacatc atatatgcct tgagcaactg 17340
atagctgtcg ctgtcaactg tcactgtaat acgctgcttc atagcacacc tctttttgac 17400
atacttctgt tcttgatgca gatgattttc aggactatga cactagcgta tatgaatagg 17460
tagatgtttt tattttgtca cacaaaaaag aggctcgcac ctctttttct tatttctttt 17520
tatgatttaa tacggcattg aggacaatag cgagtaggct ggatacgacg attccgtttg 17580

agaagaacat ttggaaggct gtcggtcgac taagttggca gcatcacccg aagaacattt 17640
ggaaggctgt cggtcgacta caggtcacta ataccatcta agtagttgat tcatagtgac 17700
tggatatgtt gtgttttaca gtattatgta gtctgttttt tatgcaaaat ctaatttaat 17760
atattgatat ttatatcatt ttacgtttct cgttcagctt ttttgtacaa agttggcatt 17820
ataaaaaagc attgctcatc aatttgttgc aacgaacagg tcactatcag tcaaaataaa 17880
atcattattt ggggcccgag atccatgcta gctctagagt cctgctttaa tgagatatgc 17940
gagacgccta tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt gtaaaaaacc 18000
tgagcatgtg tagctcagat ccttaccgcc ggtttcggtt cattctaatg aatatatcac 18060
ccgttactat cgtattttta tgaataatat tctccgttca atttactgat tgtaccctac 18120
tacttatatg tacaatatta aaatgaaaac aatatattgt gctgaatagg tttatagcga 18180
catctatgat agagcgccac aataacaaac aattgcgttt tattattaca aatccaattt 18240
taaaaaaagc ggcagaaccg gtcaaaccta aaagactgat tacataaatc ttattcaaat 18300
ttcaaaaggc cccaggggct agtatctacg acacaccgag cggcgaacta ataacgttca 18360
ctgaagggaa ctccggttcc ccgccggcgc gcatgggtga gattccttga agttgagtat 18420
tggccgtccg ctctaccgaa agttacgggc accattcaac ccggtccagc acggcggccg 18480
ggtaaccgac ttgctgcccc gagaattatg cagcattttt ttggtgtatg tgggccccaa 18540
atgaagtgca ggtcaaacct tgacagtgac gacaaatcgt tgggcgggtc cagggcgaat 18600
tttgcgacaa catgtcgagg ctcagcagga cctgcaggca tgcaagctag cttactagtg 18660
atgcatattc tatagtgtca cctaaatctg c 18691

<210> 14
<211> 59
<212> DNA
<213> Artificial sequence

<220>
<223> forward primer used for the amplification of 200 and 400 bp CHS ragments

<400> 14
ggggacaagt ttgtacaaaa aagcaggctg cactgctaac cctgagaacc atgtgcttc 59

<210> 15
<211> 59
<212> DNA
<213> Artificial sequence

<220>
<223> reverse primer for amplification of 400 bp CHS fragment

<400> 15
ggggaccact ttgtacaaga aagctgggtc gcttgacgga aggacggaga ccaagaagc 59

<210> 16
<211> 59
<212> DNA
<213> Artificial sequence

<220>
<223> reverse primer for amplification of 200bp CHS fragment

<400> 16
ggggaccact ttgtacaaga aagctgggta ggagccatgt aagcacacat gtgtgggtt 59

<210> 17
<211> 100
<212> DNA
<213> Artificial sequence

<220>
<223> forward primer for amplification of 100bp CHS fragment

<400> 17
ggggacaagt ttgtacaaaa aagcaggctg cactgctaac cctgagaacc atgtgcttca 60
ggcggagtat cctgactact acttccgcat caccaacagt 100

<210> 18
<211> 100
<212> DNA
<213> Artificial sequence

<220>
<223> reverse primer for amplification of 100 bp CHS fragment

<400> 18
ggggaccact ttgtacaaga aagctgggta acttctcctt gaggtcggtc atgtgttcac 60
tgttggtgat gcggaagtag tagtcaggat actccgcctg 100

<210> 19
<211> 79
<212> DNA
<213> Artificial sequence

<220>
<223> forward primer for amplification of 50 bp CHS fragment

<400> 19
ggggacaagt ttgtacaaaa aagcaggctg cactgctaac cctgagaacc atgtgcttca 60
ggcggagtat cctgactac 79

<210> 20
<211> 79
<212> DNA
<213> Artificial sequence

<220>
<223> reverse primer for 50 bp CHS fragment

<400> 20
ggggaccact ttgtacaaga aagctgggtg tagtcaggat actccgcctg aagcacatgg 60
ttctcagggt tagcagtgc 79

<210> 21
<211> 54
<212> DNA
<213> Artificial sequence

<220>
<223> forward primer for amplification of the 25 bp CHS fragment

<400> 21
ggggacaagt ttgtacaaaa aagcaggctg cactgctaac cctgagaacc atgt 54

<210> 22
<211> 54

<212> DNA
<213> Artificial sequence

<220>
<223> reverse primer for amplification of the 25 bp CHS fragment

<400> 22
ggggaccact ttgtacaaga aagctgggta catggttctc agggttagca gtgc 54

<210> 23
<211> 17862
<212> DNA
<213> Artificial sequence

<220>
<223> acceptor vector pHELLSGATE4

<400> 23
ggccgcacta gtgatatccc gcggccatgg cggccgggag catgcgacgt cgggcccaat 60
tcgccctata gtgagtcgta ttacaattca ctggccgtcg ttttacaacg tcgtgactgg 120
gaaaaccctg gcgttaccca acttaatcgc cttgcagcac atcccccttt cgccagctgg 180
cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag cctgaatggc 240
gaatggaaat tgtaaacgtt aatgggtttc tggagtttaa tgagctaagc acatacgtca 300
gaaaccatta ttgcgcgttc aaaagtcgcc taaggtcact atcagctagc aaatatttct 360
tgtcaaaaat gctccactga cgttccataa attcccctcg gtatccaatt agagtctcat 420
attcactctc aatccaaata atctgcaatg gcaattacct tatccgcaac ttctttacct 480
atttccgccc ggatccgggc aggttctccg gccgcttggg tggagaggct attcggctat 540
gactgggcac aacagacaat cggctgctct gatgccgccg tgttccggct gtcagcgcag 600
gggcgcccgg ttcttttttgt caagaccgac ctgtccggtg ccctgaatga actgcaggac 660
gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac 720
gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg gcaggatctc 780
ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc aatgcggcgg 840
ctgcatacgc ttgatccggc tacctgccca ttcgaccacc aagcgaaaca tcgcatcgag 900
cgagcacgta ctcggatgga agccggtctt gtcgatcagg atgatctgga cgaagagcat 960
caggggctcg cgccagccga actgttcgcc aggctcaagg cgcgcatgcc cgacggcgag 1020
gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc 1080
ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg 1140
ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg 1200
ctttacggta tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct tcttgacgag 1260
ttcttctgag cgggactctg gggttcgaaa tgaccgacca agcgacgccc aacctgccat 1320
cacgagattt cgattccacc gccgccttct atgaaaggtt gggcttcgga atcgttttcc 1380
gggacgccgg ctggatgatc ctccagcgcg gggatctcat gctggagttc ttcgcccacc 1440
ccgatccaac acttacgttt gcaacgtcca agagcaaata gaccacgaac gccggaaggt 1500
tgccgcagcg tgtggattgc gtctcaattc tctcttgcag gaatgcaatg atgaatatga 1560
tactgactat gaaactttga gggaatactg cctagcaccg tcacctcata acgtgcatca 1620
tgcatgccct gacaacatgg aacatcgcta tttttctgaa gaattatgct cgttggagga 1680
tgtcgcggca attgcagcta ttgccaacat cgaactaccc ctcacgcatg cattcatcaa 1740
tattattcat gcggggaaag gcaagattaa tccaactggc aaatcatcca gcgtgattgg 1800
taacttcagt tccagcgact tgattcgttt tggtgctacc cacgttttca ataaggacga 1860
gatggtggag taaagaagga gtgcgtcgaa gcagatcgtt caaacatttg gcaataaagt 1920
ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat 1980
tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag atgggttttt 2040
atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 2100
aactaggata aattatcgcg cgcggtgtca tctatgttac tagatcgaat taattccagg 2160
cggtgaaggg caatcagctg ttgcccgtct cactggtgaa aagaaaaacc accccagtac 2220
attaaaaacg tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa 2280
tatatcctgc caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca 2340
ctcgatacag gcagcccatc agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg 2400
cagactttgc tcatgttacc gatgctattc ggaagaacgg caactaagct gccgggtttg 2460
aaacacggat gatctcgcgg agggtagcat gttgattgta acgatgacag agcgttgctg 2520
cctgtgatca aatatcatct ccctcgcaga gatccgaatt atcagccttc ttattcattt 2580
ctcgcttaac cgtgacaggc tgtcgatctt gagaactatg ccgacataat aggaaatcgc 2640

```
tggataaagc cgctgaggaa gctgagtggc gctatttctt tagaagtgaa cgttgacgat   2700
gtcgacggat cttttccgct gcataaccct gcttcggggt cattatagcg atttttcgg    2760
tatatccatc ctttttcgca cgatatacag gattttgcca aagggttcgt gtagactttc   2820
cttggtgtat ccaacggcgt cagccgggca ggataggtga agtaggccca cccgcgagcg   2880
ggtgttcctt cttcactgtc ccttattcgc acctggcggt gctcaacggg aatcctgctc   2940
tgcgaggctg gccggctacc gccggcgtaa cagatgaggg caagcggatg gctgatgaaa   3000
ccaagccaac caggggtgat gctgccaact tactgattta gtgtatgatg gtgtttttga   3060
ggtgctccag tggcttctgt ttctatcagc tgtcctcct gttcagctac tgacggggtg    3120
gtgcgtaacg gcaaaagcac cgccggacat cagcgctatc tctgctctca ctgccgtaaa   3180
acatggcaac tgcagttcac ttacaccgct tctcaacccg gtacgcacca gaaaatcatt   3240
gatatggcca tgaatggcgt tggatgccgg gcaacagccc gcattatggg cgttggcctc   3300
aacacgattt tacgtcactt aaaaaactca ggccgcagtc ggtaacctcg cgcatacagc   3360
cgggcagtga cgtcatcgtc tgcgcggaaa tggacgaaca gtggggctat gtcggggcta   3420
aatcgcgcca gcgctggctg ttttacgcgt atgacagtct ccggaagacg gttgttgcgc   3480
acgtattcgg tgaacgcact atggcgacgc tggggcgtct tatgagcctg ctgtcaccct   3540
ttgacgtggt gatatggatg acggatggct ggccgctgta tgaatcccgc ctgaagggaa   3600
agctgcacgt aatcagcaag cgatatacgc agcgaattga gcggcataac ctgaatctga   3660
ggcagcacct ggcacggctg ggacggaagt cgctgtcgtt ctcaaaatcg gtggagctgc   3720
atgacaaagt catcgggcat tatctgaaca taaaacacta tcaataagtt ggagtcatta   3780
cccaaccagg aagggcagcc cacctatcaa ggtgtactgc cttccagacg aacgaagagc   3840
gattgaggaa aaggcggcgg cggccggcat gagcctgtcg gcctacctgc tggccgtcgg   3900
ccagggctac aaaatcacgg gcgtcgtgga ctatgagcac gtccgcgagc tggcccgcat   3960
caatggcgac ctgggccgcc tgggcggcct gctgaaactc tggctcaccg acgacccgcg   4020
cacggcgcgg ttcggtgatg ccacgatcct cgccctgctg gcgaagatcg aagagaagca   4080
ggacgagctt ggcaaggtca tgatgggcgt ggtccgcccg agggcagagc catgactttt   4140
ttagccgcta aaacggccgg ggggtgcgcg tgattgccaa gcacgtcccc atgcgctcca   4200
tcaagaagag cgacttcgcg gagctggtat tcgtgcaggg caagattcgg aataccaagt   4260
acgagaagga cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc   4320
tggacaccaa ggcaccaggc gggtcaaatc aggaataagg gcacattgcc ccggcgtgag   4380
tcggggcaat cccgcaagga gggtgaatga atcggacgtt tgaccggaag gcatacaggc   4440
aagaactgat cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg   4500
tcatgcgtgc gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg   4560
ccaagatcga gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg   4620
ccgtggagcg ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag tcgatgacca   4680
tcgacacgcg aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa   4740
aacaggtcag cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg   4800
aaatgcagct ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa   4860
acgacacggc ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc   4920
tgcaaaacaa ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg   4980
agctgcgggc cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca   5040
cccctatcgg cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt   5100
cgatcaatgg ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg   5160
cgatgggctt cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct   5220
tccgcgtcct ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa   5280
tcgtcgtgct gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc   5340
tgtcgccgac ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc   5400
tcaagctgga aaccttccgc ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc   5460
gcgagcaggt cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct   5520
gggtcaatga tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg   5580
ggggttcagc agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca   5640
cttgcttcgc tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag   5700
aggattaaaa ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt   5760
gcaggatttc cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc   5820
cgtttacgag cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc   5880
cgtggcattc ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga   5940
ggacggcccc aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca   6000
gcgaggccga ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt   6060
gatgatcgtc cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc   6120
acttaatatt tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg   6180
ggtcgcggcg acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct   6240
gctaggtagc ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt   6300
ggcgctgttg gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct   6360
ggcgggggcg gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt   6420
```

gcctctgctc acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt 6480
agctttagtg tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc 6540
gtggctcggc ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg 6600
actcgaacct acagttgttt ccttactggg ctttctcagc cgggatggcg ctaagaagct 6660
attgccgccg atcttcatat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac 6720
cgcatcaggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg 6780
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat 6840
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc 6900
gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa aaatcgacgc 6960
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga 7020
agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt 7080
ctcccttcgg gaagcgtggc gctttctcaa tgctcacgct gtaggtatct cagttcggtg 7140
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc 7200
gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg 7260
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc 7320
ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg 7380
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc 7440
gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatat 7500
caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt 7560
taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa 7620
aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa 7680
tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc 7740
tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct 7800
gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca 7860
gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt 7920
aaacaagtgg cagcaacgga ttcgcaaacc tgtcacgcct tttgtgccaa aagccgcgcc 7980
aggtttgcga tccgctgtgc caggcgttag gcgtcatatg aagatttcgg tgatccctga 8040
gcaggtggcg gaaacattgg atgctgagaa ccatttcatt gttcgtgaag tgttcgatgt 8100
gcacctatcc gaccaaggct ttgaactatc taccagaagt gtgagcccct accggaagga 8160
ttacatctcg gatgatgact ctgatgaaga ctctgcttgc tatggcgcat tcatcgacca 8220
agagcttgtc gggaagattg aactcaactc aacatggaac gatctagcct ctatcgaaca 8280
cattgttgtg tcgcacacgc accgaggcaa aggagtcgcg cacagtctca tcgaatttgc 8340
gaaaaagtgg gcactaagca gacagctcct tggcatacga ttagagacac aaacgaacaa 8400
tgtacctgcc tgcaatttgt acgcaaaatg tggctttact ctcggcggca ttgacctgtt 8460
cacgtataaa actagacctc aagtctcgaa cgaaacagcg atgtactggt actggttctc 8520
gggagcacag gatgacgcct aacaattcat tcaagccgac accgcttcgc ggcgcggctt 8580
aattcaggag ttaaacatca tgagggaagc ggtgatcgcc gaagtatcga ctcaactatc 8640
agaggtagtt ggcgtcatcg agcgccatct cgaaccgacg ttgctggccg tacatttgta 8700
cggctccgca gtggatggcg gcctgaagcc acacagtgat attgatttgc tggttacggt 8760
gaccgtaagg cttgatgaaa caacgcggcg agctttgatc aacgaccttt tggaaacttc 8820
ggcttcccct ggagagagcg agattctccg cgctgtagaa gtcaccattg ttgtgcacga 8880
cgacatcatt ccgtggcgtt atccagctaa gcgcgaactg caatttggag aatggcagcg 8940
caatgacatt cttgcaggta tcttcgagcc agccacgatc gacattgatc tggctatctt 9000
gctgacaaaa gcaagagaac atagcgttgc cttggtaggt ccagcggcgg aggaactctt 9060
tgatccggtt cctgaacagg atctatttga ggcgctaaat gaaaccttaa cgctatggaa 9120
ctcgccgccc gactgggctg gcgatgagcg aaatgtagtg cttacgttgt cccgcatttg 9180
gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc gctgccgact gggcaatgga 9240
gcgcctgccg gcccagtatc agcccgtcat acttgaagct aggcaggctt atcttggaca 9300
agaagatcgc ttggcctcgc gcgcagatca gttggaagaa tttgttcact acgtgaaagg 9360
cgagatcacc aaggtagtcg gcaaataatg tctaacaatt cgttcaagcc gacgccgctt 9420
cgcggcgcgg cttaactcaa gcgttagaga gctggggaag actatgcgcg atctgttgaa 9480
ggtggttcta agcctcgtac ttgcgatggc atcggggcag gcacttgctg acctgccaat 9540
tgttttagtg gatgaagctc gtcttcccta tgactactcc ccatccaact acgacatttc 9600
tccaagcaac tacgacaact ccataagcaa ttacgacaat agtccatcaa attacgacaa 9660
ctctgagagc aactacgata atagttcatc caattacgac aatagtcgca acggaaatcg 9720
taggcttata tatagcgcaa atgggtctcg cactttcgcc ggctactacg tcattgccaa 9780
caatgggaca acgaacttct tttccacatc tggcaaaagg atgttctaca ccccaaaagg 9840
ggggcgcggc gtctatggcg gcaaagatgg gagcttctgc ggggcattgg tcgtcataaa 9900
tggccaattt tcgcttgccc tgacagataa cggcctgaag atcatgtatc taagcaacta 9960
gcctgctctc taataaaatg ttaggagctt ggctgccatt tttggggtga ggccgttcgc 10020
ggccgagggg cgcagcccct ggggggatgg gaggcccgcg ttagcgggcc gggagggttc 10080
gagaaggggg ggcacccccc ttcggcgtgc gcggtcacgc gccagggcgc agccctggtt 10140
aaaaacaagg tttataaata ttggtttaaa agcaggttaa aagacaggtt agcggtggcc 10200

```
gaaaaacggg cggaaaccct tgcaaatgct ggattttctg cctgtggaca gcccctcaaa  10260
tgtcaatagg tgcgcccctc atctgtcagc actctgcccc tcaagtgtca aggatcgcgc  10320
ccctcatctg tcagtagtcg cgcccctcaa gtgtcaatac cgcagggcac ttatccccag  10380
gcttgtccac atcatctgtg ggaaactcgc gtaaaatcag gcgttttcgc cgatttgcga  10440
ggctggccag ctccacgtcg ccggccgaaa tcgagcctgc ccctcatctg tcaacgccgc  10500
gccgggtgag tcggcccctc aagtgtcaac gtccgccccc catctgtcag tgagggccaa  10560
gttttccgcg aggtatccac aacgccggcg gccggccgcg gtgtctcgca cacggcttcg  10620
acggcgtttc tggcgcgttt gcagggccat agacggccgc cagcccagcg gcgagggcaa  10680
ccagcccggt gagcgtcgga aagggtcgac atcttgctgc gttcggatat tttcgtggag  10740
ttcccgccac agacccggat tgaaggcgag atccagcaac tcgcgccaga tcatcctgtg  10800
acggaacttt ggcgcgtgat gactggccag gacgtcggcc gaaagagcga caagcagatc  10860
acgattttcg acagcgtcgg atttgcgatc gaggattttt cggcgctgcg ctacgtccgc  10920
gaccgcgttg agggatcaag ccacagcagc ccactcgacc ttctagccga cccagacgag  10980
ccaagggatc tttttggaat gctgctccgt cgtcaggctt tccgacgttt gggtggttga  11040
acagaagtca ttatcgtacg gaatgccagc actcccgagg ggaaccctgt ggttggcatg  11100
cacatacaaa tggacgaacg gataaacctt ttcacgccct tttaaatatc cgttattcta  11160
ataaacgctc ttttctctta ggtttacccg ccaatatatc ctgtcaaaca ctgatagttt  11220
aaactgaagg cgggaaacga caatctgatc atgagcggag aattaaggga gtcacgttat  11280
gacccccgcc gatgacgcgg gacaagccgt tttacgtttg gaactgacag aaccgcaacg  11340
attgaaggag ccactcagcc ccaatacgca aaccgcctct ccccgcgcgt tggccgattc  11400
attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa  11460
ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg cttccggctc  11520
gtatgttgtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg  11580
attacgccaa gctatttagg tgacactata gaatactcaa gctatgcatc caacgcgttg  11640
ggagctctcc catatcgacc tgcaggcggc cgctcgacga attaattcca atcccacaaa  11700
aatctgagct taacagcaca gttgctcctc tcagagcaga atcgggtatt caacaccctc  11760
atatcaacta ctacgttgtg tataacggtc cacatgccgg tatatacgat gactggggtt  11820
gtacaaaggc ggcaacaaac ggcgttcccg gagttgcaca caagaaattt gccactatta  11880
cagaggcaag agcagcagct gacgcgtaca caacaagtca gcaaacagac aggttgaact  11940
tcatccccaa aggagaagct caactcaagc ccaagagctt tgctaaggcc ctaacaagcc  12000
caccaaagca aaaagcccac tggctcacgc taggaaccaa aaggcccagc agtgatccag  12060
ccccaaaaga gatctccttt gccccggaga ttacaatgga cgatttcctc tatctttacg  12120
atctaggaag gaagttcgaa ggtgaaggtg acgacactat gttcaccact gataatgaga  12180
aggttagcct cttcaatttc agaaagaatg ctgacccaca gatggttaga gaggcctacg  12240
cagcaggtct catcaagacg atctacccga gtaacaatct ccaggagatc aaataccttc  12300
ccaagaaggt taaagatgca gtcaaaagat tcaggactaa ttgcatcaag aacacagaga  12360
aagacatatt tctcaagatc agaagtacta ttccagtatg gacgattcaa ggcttgcttc  12420
ataaaccaag gcaagtaata gagattggag tctctaaaaa ggtagttcct actgaatcta  12480
aggccatgca tggagtctaa gattcaaatc gaggatctaa cagaactcgc cgtgaagact  12540
ggcgaacagt tcatacagag tcttttacga ctcaatgaca agaagaaaat cttcgtcaac  12600
atggtggagc acgacactct ggtctactcc aaaaatgtca aagatacagt ctcagaagac  12660
caaagggcta ttgagacttt tcaacaaagg ataatttcgg gaaacctcct cggattccat  12720
tgcccagcta tctgtcactt catcgaaagg acagtagaaa aggaaggtgg ctcctacaaa  12780
tgccatcatt gcgataaagg aaaggctatc attcaagatc tctctgccga cagtggtccc  12840
aaagatggac ccccacccac gaggagcatc gtggaaaaag aagacgttcc aaccacgtct  12900
tcaaagcaag tggattgatg tgacatctcc actgacgtaa gggatgacgc acaatcccac  12960
tatccttcgc aagacccttc ctctatataa ggaagttcat ttcatttgga gaggacacgc  13020
tcgaggctag catggatctc gggccccaaa taatgatttt attttgactg atagtgacct  13080
gttcgttgca acaaattgat gagcaatgct tttttataat gccaactttg tacaaaaaag  13140
ctgaacgaga aacgtaaaat gatataaata tcaatatatt aaattagatt ttgcataaaa  13200
aacagactac ataatactgt aaaacacaac atatccagtc actatgaatc aactacttag  13260
atggtattag tgacctgtag tcgaccgaca gccttccaaa tgttcttcgg gtgatgctgc  13320
caacttagtc gaccgacagc cttccaaatg ttcttctcaa acggaatcgt cgtatccagc  13380
ctactcgcta ttgtcctcaa tgccgtatta aatcataaaa agaaataaga aaaagaggtg  13440
cgagcctctt ttttgtgtga caaaataaaa acatctacct attcatatac gctagtgtca  13500
tagtcctgaa aatcatctgc atcaagaaca atttcacaac tcttatactt ttctcttaca  13560
agtcgttcgg cttcatctgg attttcagcc tctatactta ctaaacgtga taaagtttct  13620
gtaatttcta ctgtatcgac ctgcagactg gctgtgtata agggagcctg acatttatat  13680
tccccagaac atcaggttaa tggcgttttt gatgtcattt tcgcggtggc tgagatcagc  13740
cacttcttcc ccgataacgg agaccggcac actggccata tcggtggtca tcatgcgcca  13800
gctttcatcc ccgatatgca ccaccgggta aagttcacgg gagactttat ctgacagcag  13860
acgtgcactg gccaggggga tcaccatccg tcgcccgggc gtgtcaataa tatcactctg  13920
tacatccaca aacagacgat aacggctctc tcttttatag gtgtaaacct taaactgcat  13980
```

```
ttcaccagtc cctgttctcg tcagcaaaag agccgttcat ttcaataaac cgggcgacct  14040
cagccatccc ttcctgattt tccgctttcc agcgttcggc acgcagacga cgggcttcat  14100
tctgcatggt tgtgcttacc agaccggaga tattgacatc atatatgcct tgagcaactg  14160
atagctgtcg ctgtcaactg tcactgtaat acgctgcttc atagcacacc tctttttgac  14220
atacttcggg tagtgccgat caacgtctca ttttcgccaa aagttggccc agggcttccc  14280
ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg tcacaggtat  14340
ttattcggcg caaagtgcgt cgggtgatgc tgccaactta gtcgactaca ggtcactaat  14400
accatctaag tagttgattc atagtgactg gatatgttgt gttttacagt attatgtagt  14460
ctgtttttta tgcaaaatct aatttaatat attgatattt atatcatttt acgtttctcg  14520
ttcagctttc ttgtacaaag ttggcattat aagaaagcat tgcttatcaa tttgttgcaa  14580
cgaacaggtc actatcagtc aaaataaaat cattatttgc catccagctg cagctcctcg  14640
aggaattcgg taccccagct tggtaaggaa ataattattt tctttttcc ttttagtata  14700
aaatagttaa gtgatgttaa ttagtatgat tataataata tagttgttat aattgtgaaa  14760
aaataattta taaatatatt gtttacataa acaacatagt aatgtaaaaa aatatgacaa  14820
gtgatgtgta agacgaagaa gataaaagtt gagagtaagt atattatttt taatgaattt  14880
gatcgaacat gtaagatgat atactagcat taatatttgt tttaatcata atagtaattc  14940
tagctggttt gatgaattaa atatcaatga taaaatacta tagtaaaaat aagaataaat  15000
aaattaaaat aatatttttt tatgattaat agtttattat ataattaaat atctatacca  15060
ttactaaata ttttagttta aaagttaata aatattttgt tagaaattcc aatctgcttg  15120
taatttatca ataaacaaaa tattaaataa caagctaaag taacaaataa tatcaaacta  15180
atagaaacag taatctaatg taacaaaaca taatctaatg ctaatataac aaagcgcaag  15240
atctatcatt ttatatagta ttattttcaa tcaacattct tattaatttc taaataatac  15300
ttgtagtttt attaacttct aaatggattg actattaatt aaatgaatta gtcgaacatg  15360
aataaacaag gtaacatgat agatcatgtc attgtgttat cattgatctt acatttggat  15420
tgattacagt tgggaagctg ggttcgaaat cgataagctt ggatcctcta gagagctgca  15480
gctggatggc aaataatgat tttattttga ctgatagtga cctgttcgtt gcaacaaatt  15540
gataagcaat gctttcttat aatgccaact ttgtacaaga aagctgaacg agaaacgtaa  15600
aatgatataa atatcaatat attaaattag attttgcata aaaaacagac tacataatac  15660
tgtaaaacac aacatatcca gtcactatga atcaactact tagatggtat tagtgacctg  15720
tagtcgacta agttggcagc atcacccgac gcactttgcg ccgaataaat acctgtgacg  15780
gaagatcact tcgcagaata ataaatcct ggtgtccctg ttgataccgg gaagccctgg  15840
gccaactttt ggcgaaaatg agacgttgat cggcactacc catttcacaa ctcttatact  15900
tttctcttac aagtcgttcg gcttcatctg gattttcagc ctctatactt actaaacgtg  15960
ataaagtttc tgtaatttct actgtatcga cctgcagact ggctgtgtat aagggagcct  16020
gacatttata ttccccagaa catcaggtta atggcgtttt tgatgtcatt ttcgcggtgg  16080
ctgagatcag ccacttcttc cccgataacg gagaccggca cactggccat atcggtggtc  16140
atcatgcgcc agctttcatc cccgatatgc accaccgggt aaagttcacg ggagacttta  16200
tctgacagca gacgtgcact ggccaggggg atcaccatcc gtcgcccggg cgtgtcaata  16260
atatcactct gtacatccac aaacagacga taacggctct ctcttttata ggtgtaaacc  16320
ttaaactgca tttcaccagt ccctgttctc gtcagcaaaa gagccgttca tttcaataaa  16380
ccgggcgacc tcagccatcc cttcctgatt ttccgctttc cagcgttcgg cacgcagacg  16440
acgggcttca ttctgcatgg ttgtgcttac cagaccggag atattgacat catatatgcc  16500
ttgagcaact gatagctgtc gctgtcaact gtcactgtaa tacgctgctt catagcacac  16560
ctctttttga catacttctg ttcttgatgc agatgatttt caggactatg acactagcgt  16620
atatgaatag gtagatgttt ttattttgtc acacaaaaaa gaggctcgca cctctttttc  16680
ttatttcttt ttatgattta atacggcatt gaggacaata gcgagtaggc tggatacgac  16740
gattccgttt gagaagaaca tttggaaggc tgtcggtcga ctaagttggc agcatcaccc  16800
gaagaacatt tggaaggctg tcggtcgact acaggtcact aataccatct aagtagttga  16860
ttcatagtga ctggatatgt tgtgttttac agtattatgt agtctgtttt ttatgcaaaa  16920
tctaatttaa tatattgata tttatatcat tttacgtttc tcgttcagct tttttgtaca  16980
aagttggcat tataaaaaag cattgctcat caatttgttg caacgaacag gtcactatca  17040
gtcaaaataa aatcattatt tggggcccga gatccatgct agctctagag tcctgcttta  17100
atgagatatg cgagacgcct atgatcgcat gatatttgct ttcaattctg ttgtgcacgt  17160
tgtaaaaaac ctgagcatgt gtagctcaga tccttaccgc cggtttcggt tcattctaat  17220
gaatatatca cccgttacta tcgtatttt atgaataata ttctccgttc aatttactga  17280
ttgtacccta ctacttatat gtacaatatt aaaatgaaaa caatatattg tgctgaatag  17340
gtttatagcg acatctatga tagagcgcca caataacaaa caattgcgtt ttattattac  17400
aaatccaatt ttaaaaaaag cggcagaacc ggtcaaacct aaaagactga ttacataaat  17460
cttattcaaa tttcaaaagg ccccaggggc tagtatctac gacacaccga gcggcgaact  17520
aataacgttc actgaaggga actccggttc cccgccggcg cgcatgggtg agattccttg  17580
aagttgagta ttggccgtcc gctctaccga aagttacggg caccattcaa cccggtccag  17640
cacggcggcc gggtaaccga cttgctgccc cgagaattat gcagcatttt tttggtgtat  17700
gtgggcccca aatgaagtgc aggtcaaacc ttgacagtga cgacaaatcg ttgggcgggt  17760
```

ccagggcgaa ttttgcgaca acatgtcgag gctcagcagg acctgcaggc atgcaagcta 17820
gcttactagt gatgcatatt ctatagtgtc acctaaatct gc 17862

<210> 24
<211> 17476
<212> DNA
<213> Artificial sequence

<220>
<223> acceptor vector pHELLSGATE8

<400> 24
ggccgcacta gtgatatccc gcggccatgg cggccgggag catgcgacgt cgggcccaat 60
tcgccctata gtgagtcgta ttacaattca ctggccgtcg ttttacaacg tcgtgactgg 120
gaaaaccctg gcgttaccca acttaatcgc cttgcagcac atcccccttt cgccagctgg 180
cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag cctgaatggc 240
gaatggaaat tgtaaacgtt aatgggtttc tggagtttaa tgagctaagc acatacgtca 300
gaaaccatta ttgcgcgttc aaaagtcgcc taaggtcact atcagctagc aaatatttct 360
tgtcaaaaat gctccactga cgttccataa attcccctcg gtatccaatt agagtctcat 420
attcactctc aatccaaata atctgcaatg gcaattacct tatccgcaac ttctttacct 480
atttccgccc ggatccgggc aggttctccg gccgcttggg tggagaggct attcggctat 540
gactgggcac aacagacaat cggctgctct gatgccgccg tgttccggct gtcagcgcag 600
gggcgcccgg ttctttttgt caagaccgac ctgtccggtg ccctgaatga actgcaggac 660
gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac 720
gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg gcaggatctc 780
ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc aatgcggcgg 840
ctgcatacgc ttgatccggc tacctgccca ttcgaccacc aagcgaaaca tcgcatcgag 900
cgagcacgta ctcggatgga agccggtctt gtcgatcagg atgatctgga cgaagagcat 960
caggggctcg cgccagccga actgttcgcc aggctcaagg cgcgcatgcc cgacggcgag 1020
gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc 1080
ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg 1140
ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg 1200
ctttacggta tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct tcttgacgag 1260
ttcttctgag cgggactctg gggttcgaaa tgaccgacca agcgacgccc aacctgccat 1320
cacgagattt cgattccacc gccgccttct atgaaaggtt gggcttcgga atcgttttcc 1380
gggacgccgg ctggatgatc ctccagcgcg gggatctcat gctggagttc ttcgcccacc 1440
ccgatccaac acttacgttt gcaacgtcca agagcaaata gaccacgaac gccggaaggt 1500
tgccgcagcg tgtggattgc gtctcaattc tctcttgcag gaatgcaatg atgaatatga 1560
tactgactat gaaactttga gggaatactg cctagcaccg tcacctcata acgtgcatca 1620
tgcatgccct gacaacatgg aacatcgcta tttttctgaa gaattatgct cgttggagga 1680
tgtcgcggca attgcagcta ttgccaacat cgaactaccc ctcacgcatg cattcatcaa 1740
tattattcat gcggggaaag gcaagattaa tccaactggc aaatcatcca gcgtgattgg 1800
taacttcagt tccagcgact tgattcgttt tggtgctacc cacgttttca ataaggacga 1860
gatggtggag taaagaagga gtgcgtcgaa gcagatcgtt caaacatttg gcaataaagt 1920
ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat 1980
tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag atgggttttt 2040
atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 2100
aactaggata aattatcgcg cgcggtgtca tctatgttac tagatcgaat taattccagg 2160
cggtgaaggg caatcagctg ttgcccgtct cactggtgaa aagaaaaacc accccagtac 2220
attaaaaacg tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa 2280
tatatcctgc caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca 2340
ctcgatacag gcagcccatc agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg 2400
cagactttgc tcatgttacc gatgctattc ggaagaacgg caactaagct gccgggtttg 2460
aaacacggat gatctcgcgg agggtagcat gttgattgta acgatgacag agcgttgctg 2520
cctgtgatca aatatcatct ccctcgcaga gatccgaatt atcagccttc ttattcattt 2580
ctcgcttaac cgtgacaggc tgtcgatctt gagaactatg ccgacataat aggaaatcgc 2640
tggataaagc cgctgaggaa gctgagtggc gctatttctt tagaagtgaa cgttgacgat 2700
gtcgacggat cttttccgct gcataaccct gcttcggggt cattatagcg atttttttcgg 2760
tatatccatc ctttttcgca cgatatacag gattttgcca aagggttcgt gtagactttc 2820
cttggtgtat ccaacggcgt cagccgggca ggataggtga agtaggccca cccgcgagcg 2880
ggtgttcctt cttcactgtc ccttattcgc acctggcggt gctcaacggg aatcctgctc 2940
tgcgaggctg gccggctacc gccggcgtaa cagatgaggg caagcggatg gctgatgaaa 3000

```
ccaagccaac caggggtgat gctgccaact tactgattta gtgtatgatg gtgtttttga   3060
ggtgctccag tggcttctgt ttctatcagc tgtccctcct gttcagctac tgacggggtg   3120
gtgcgtaacg gcaaaagcac cgccggacat cagcgctatc tctgctctca ctgccgtaaa   3180
acatggcaac tgcagttcac ttacaccgct tctcaacccg gtacgcacca gaaaatcatt   3240
gatatggcca tgaatggcgt tggatgccgg gcaacagccc gcattatggg cgttggcctc   3300
aacacgattt tacgtcactt aaaaaactca ggccgcagtc ggtaacctcg cgcatacagc   3360
cgggcagtga cgtcatcgtc tgcgcggaaa tggacgaaca gtggggctat gtcggggcta   3420
aatcgcgcca gcgctggctg ttttacgcgt atgacagtct ccggaagacg gttgttgcgc   3480
acgtattcgg tgaacgcact atggcgacgc tggggcgtct tatgagcctg ctgtcaccct   3540
ttgacgtggt gatatggatg acggatggct ggccgctgta tgaatcccgc ctgaagggaa   3600
agctgcacgt aatcagcaag cgatatacgc agcgaattga gcggcataac ctgaatctga   3660
ggcagcacct ggcacggctg ggacggaagt cgctgtcgtt ctcaaaatcg gtggagctgc   3720
atgacaaagt catcgggcat tatctgaaca taaaacacta tcaataagtt ggagtcatta   3780
cccaaccagg aagggcagcc cacctatcaa ggtgtactgc cttccagacg aacgaagagc   3840
gattgaggaa aaggcggcgg cggccggcat gagcctgtcg gcctacctgc tggccgtcgg   3900
ccagggctac aaaatcacgg gcgtcgtgga ctatgagcac gtccgcgagc tggcccgcat   3960
caatggcgac ctgggccgcc tgggcggcct gctgaaactc tggctcaccg acgacccgcg   4020
cacggcgcgg ttcggtgatg ccacgatcct cgccctgctg gcgaagatcg aagagaagca   4080
ggacgagctt ggcaaggtca tgatgggcgt ggtccgcccg agggcagagc catgactttt   4140
ttagccgcta aaacggccgg ggggtgcgcg tgattgccaa gcacgtcccc atgcgctcca   4200
tcaagaagag cgacttcgcg gagctggtat tcgtgcaggg caagattcgg aataccaagt   4260
acgagaagga cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc   4320
tggacaccaa ggcaccaggc gggtcaaatc aggaataagg gcacattgcc ccggcgtgag   4380
tcggggcaat cccgcaagga gggtgaatga atcggacgtt tgaccggaag gcatacaggc   4440
aagaactgat cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg   4500
tcatgcgtgc gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg   4560
ccaagatcga gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg   4620
ccgtggagcg ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag tcgatgacca   4680
tcgacacgcg aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa   4740
aacaggtcag cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg   4800
aaatgcagct ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa   4860
acgacacggc ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc   4920
tgcaaaacaa ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg   4980
agctgcgggc cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca   5040
cccctatcgg cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt   5100
cgatcaatgg ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg   5160
cgatgggctt cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct   5220
tccgcgtcct ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa   5280
tcgtcgtgct gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc   5340
tgtcgccgac ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc   5400
tcaagctgga aaccttccgc ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc   5460
gcgagcaggt cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct   5520
gggtcaatga tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg   5580
ggggttcagc agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca   5640
cttgcttcgc tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag   5700
aggattaaaa ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt   5760
gcaggatttc cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc   5820
cgtttacgag cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc   5880
cgtggcattc ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga   5940
ggacggcccc aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca   6000
gcgaggccga ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt   6060
gatgatcgtc cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc   6120
acttaatatt tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg   6180
ggtcgcggcg acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct   6240
gctaggtagc ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt   6300
ggcgctgttg gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct   6360
ggcgggggcg gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt   6420
gcctctgctc acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt   6480
agctttagtg tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc   6540
gtggctcggc ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg   6600
actcgaacct acagttgttt ccttactggg ctttctcagc cgggatggcg ctaagaagct   6660
attgccgccg atcttcatat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac   6720
cgcatcaggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg   6780
```

```
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat   6840
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc   6900
gcgttgctgg cgtttttcca taggctccgc ccccctgacg agcatcacaa aaatcgacgc   6960
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga   7020
agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt   7080
ctcccttcgg gaagcgtggc gctttctcaa tgctcacgct gtaggtatct cagttcggtg   7140
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc   7200
gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg   7260
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc   7320
ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg   7380
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc   7440
gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatat   7500
caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt   7560
taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa   7620
aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa   7680
tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc   7740
tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct   7800
gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca   7860
gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt   7920
aaacaagtgg cagcaacgga ttcgcaaacc tgtcacgcct tttgtgccaa aagccgcgcc   7980
aggtttgcga tccgctgtgc caggcgttag gcgtcatatg aagatttcgg tgatccctga   8040
gcaggtggcg gaaacattgg atgctgagaa ccatttcatt gttcgtgaag tgttcgatgt   8100
gcacctatcc gaccaaggct ttgaactatc taccagaagt gtgagccccc accggaagga   8160
ttacatctcg gatgatgact ctgatgaaga ctctgcttgc tatggcgcat tcatcgacca   8220
agagcttgtc gggaagattg aactcaactc aacatggaac gatctagcct ctatcgaaca   8280
cattgttgtg tcgcacacgc accgaggcaa aggagtcgcg cacagtctca tcgaatttgc   8340
gaaaaagtgg gcactaagca gacagctcct tggcatacga ttagagacac aaacgaacaa   8400
tgtacctgcc tgcaatttgt acgcaaaatg tggctttact ctcggcggca ttgacctgtt   8460
cacgtataaa actagacctc aagtctcgaa cgaaacagcg atgtactggt actggttctc   8520
gggagcacag gatgacgcct aacaattcat tcaagccgac accgcttcgc ggcgcggctt   8580
aattcaggag ttaaacatca tgagggaagc ggtgatcgcc gaagtatcga ctcaactatc   8640
agaggtagtt ggcgtcatcg agcgccatct cgaaccgacg ttgctggccg tacatttgta   8700
cggctccgca gtggatggcg gcctgaagcc acacagtgat attgatttgc tggttacggt   8760
gaccgtaagg cttgatgaaa caacgcggcg agctttgatc aacgaccttt tggaaacttc   8820
ggcttcccct ggagagagcg agattctccg cgctgtagaa gtcaccattg ttgtgcacga   8880
cgacatcatt ccgtggcgtt atccagctaa gcgcgaactg caatttggag aatggcagcg   8940
caatgacatt cttgcaggta tcttcgagcc agccacgatc gacattgatc tggctatctt   9000
gctgacaaaa gcaagagaac atagcgttgc cttggtaggt ccagcggcgg aggaactctt   9060
tgatccggtt cctgaacagg atctatttga ggcgctaaat gaaaccttaa cgctatggaa   9120
ctcgccgccc gactgggctg gcgatgagcg aaatgtagtg cttacgttgt cccgcatttg   9180
gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc gctgccgact gggcaatgga   9240
gcgcctgccg gcccagtatc agcccgtcat acttgaagct aggcaggctt atcttggaca   9300
agaagatcgc ttggcctcgc gcgcagatca gttggaagaa tttgttcact acgtgaaagg   9360
cgagatcacc aaggtagtcg gcaaataatg tctaacaatt cgttcaagcc gacgccgctt   9420
cgcggcgcgg cttaactcaa gcgttagaga gctggggaag actatgcgcg atctgttgaa   9480
ggtggttcta agcctcgtac ttgcgatggc atcggggcag gcacttgctg acctgccaat   9540
tgttttagtg gatgaagctc gtcttcccta tgactactcc ccatccaact acgacatttc   9600
tccaagcaac tacgacaact ccataagcaa ttacgacaat agtccatcaa attacgacaa   9660
ctctgagagc aactacgata atagttcatc caattacgac aatagtcgca acggaaatcg   9720
taggcttata tatagcgcaa atgggtctcg cactttcgcc ggctactacg tcattgccaa   9780
caatgggaca acgaacttct tttccacatc tggcaaaagg atgttctaca ccccaaaagg   9840
ggggcgcggc gtctatggcg gcaaagatgg gagcttctgc ggggcattgg tcgtcataaa   9900
tggccaattt tcgcttgccc tgacagataa cggcctgaag atcatgtatc taagcaacta   9960
gcctgctctc taataaaatg ttaggagctt ggctgccatt tttggggtga ggccgttcgc  10020
ggccgagggg cgcagccccc ggggggatgg gaggcccgcg ttagcgggcc gggagggttc  10080
gagaaggggg ggcacccccc ttcggcgtgc gcggtcacgc gccagggcgc agccctggtt  10140
aaaaacaagg tttataaata ttggtttaaa agcaggttaa aagacaggtt agcggtggcc  10200
gaaaaacggg cggaaaccct tgcaaatgct ggattttctg cctgtggaca gcccctcaaa  10260
tgtcaatagg tgcgcccctc atctgtcagc actctgcccc tcaagtgtca aggatcgcgc  10320
ccctcatctg tcagtagtcg cgcccctcaa gtgtcaatac cgcagggcac ttatccccag  10380
gcttgtccac atcatctgtg ggaaactcgc gtaaaatcag gcgttttcgc cgatttgcga  10440
ggctggccag ctccacgtcg ccggccgaaa tcgagcctgc ccctcatctg tcaacgccgc  10500
gccgggtgag tcggcccctc aagtgtcaac gtccgccccc catctgtcag tgagggccaa  10560
```

```
gttttccgcg aggtatccac aacgccggcg gccggccgcg gtgtctcgca cacggcttcg  10620
acggcgtttc tggcgcgttt gcagggccat agacggccgc cagcccagcg gcgagggcaa  10680
ccagcccggt gagcgtcgga aagggtcgac atcttgctgc gttcggatat tttcgtggag  10740
ttcccgccac agacccggat tgaaggcgag atccagcaac tcgcgccaga tcatcctgtg  10800
acggaacttt ggcgcgtgat gactggccag gacgtcggcc gaaagagcga caagcagatc  10860
acgattttcg acagcgtcgg atttgcgatc gaggattttt cggcgctgcg ctacgtccgc  10920
gaccgcgttg agggatcaag ccacagcagc ccactcgacc ttctagccga cccagacgag  10980
ccaagggatc tttttggaat gctgctccgt cgtcaggctt tccgacgttt gggtggttga  11040
acagaagtca ttatcgtacg gaatgccagc actcccgagg ggaaccctgt ggttggcatg  11100
cacatacaaa tggacgaacg gataaacctt ttcacgccct tttaaatatc cgttattcta  11160
ataaacgctc ttttctctta ggtttacccg ccaatatatc ctgtcaaaca ctgatagttt  11220
aaactgaagg cgggaaacga caatctgatc atgagcggag aattaaggga gtcacgttat  11280
gacccccgcc gatgacgcgg gacaagccgt tttacgtttg gaactgacag aaccgcaacg  11340
attgaaggag ccactcagcc ccaatacgca aaccgcctct ccccgcgcgt tggccgattc  11400
attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa  11460
ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg cttccggctc  11520
gtatgttgtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg  11580
attacgccaa gctatttagg tgacactata gaatactcaa gctatgcatc caacgcgttg  11640
ggagctctcc catatcgacc tgcaggcggc cgctcgacga attaattcca atcccacaaa  11700
aatctgagct taacagcaca gttgctcctc tcagagcaga atcgggtatt caacaccctc  11760
atatcaacta ctacgttgtg tataacggtc cacatgccgg tatatacgat gactggggtt  11820
gtacaaaggc ggcaacaaac ggcgttcccg gagttgcaca caagaaattt gccactatta  11880
cagaggcaag agcagcagct gacgcgtaca caacaagtca gcaaacagac aggttgaact  11940
tcatccccaa aggagaagct caactcaagc ccaagagctt tgctaaggcc ctaacaagcc  12000
caccaaagca aaaagcccac tggctcacgc taggaaccaa aaggcccagc agtgatccag  12060
ccccaaaaga gatctccttt gccccggaga ttacaatgga cgatttcctc tatctttacg  12120
atctaggaag gaagttcgaa ggtgaaggtg acgacactat gttcaccact gataatgaga  12180
aggttagcct cttcaatttc agaaagaatg ctgacccaca gatggttaga gaggcctacg  12240
cagcaggtct catcaagacg atctacccga gtaacaatct ccaggagatc aaataccttc  12300
ccaagaaggt taaagatgca gtcaaaagat tcaggactaa ttgcatcaag aacacagaga  12360
aagacatatt tctcaagatc agaagtacta ttccagtatg gacgattcaa ggcttgcttc  12420
ataaaccaag gcaagtaata gagattggag tctctaaaaa ggtagttcct actgaatcta  12480
aggccatgca tggagtctaa gattcaaatc gaggatctaa cagaactcgc cgtgaagact  12540
ggcgaacagt tcatacagag tcttttacga ctcaatgaca agaagaaaat cttcgtcaac  12600
atggtggagc acgacactct ggtctactcc aaaaatgtca aagatacagt ctcagaagac  12660
caaagggcta ttgagacttt tcaacaaagg ataatttcgg gaaacctcct cggattccat  12720
tgcccagcta tctgtcactt catcgaaagg acagtagaaa aggaaggtgg ctcctacaaa  12780
tgccatcatt gcgataaagg aaaggctatc attcaagatc tctctgccga cagtggtccc  12840
aaagatggac ccccacccac gaggagcatc gtggaaaaag aagacgttcc aaccacgtct  12900
tcaaagcaag tggattgatg tgacatctcc actgacgtaa gggatgacgc acaatcccac  12960
tatccttcgc aagacccttc ctctatataa ggaagttcat ttcatttgga gaggacacgc  13020
tcgagacaag tttgtacaaa aaagctgaac gagaaacgta aaatgatata aatatcaata  13080
tattaaatta gattttgcat aaaaaacaga ctacataata ctgtaaaaca caacatatcc  13140
agtcactatg aatcaactac ttagatggta ttagtgacct gtagtcgacc gacagccttc  13200
caaatgttct tcgggtgatg ctgccaactt agtcgaccga cagccttcca aatgttcttc  13260
tcaaacggaa tcgtcgtatc cagcctactc gctattgtcc tcaatgccgt attaaatcat  13320
aaaaagaaat aagaaaaaga ggtgcgagcc tcttttttgt gtgacaaaat aaaaacatct  13380
acctattcat atacgctagt gtcatagtcc tgaaaatcat ctgcatcaag aacaatttca  13440
caactcttat acttttctct tacaagtcgt tcggcttcat ctggattttc agcctctata  13500
cttactaaac gtgataaagt ttctgtaatt tctactgtat cgacctgcag actggctgtg  13560
tataagggag cctgacattt atattcccca gaacatcagg ttaatggcgt ttttgatgtc  13620
attttcgcgg tggctgagat cagccacttc ttccccgata acggagaccg gcacactggc  13680
catatcggtg gtcatcatgc gccagctttc atccccgata tgcaccaccg ggtaaagttc  13740
acgggagact ttatctgaca gcagacgtgc actggccagg gggatcacca tccgtcgccc  13800
gggcgtgtca ataatatcac tctgtacatc cacaaacaga cgataacggc tctctctttt  13860
ataggtgtaa accttaaact gcatttcacc agtccctgtt ctcgtcagca aaagagccgt  13920
tcatttcaat aaaccgggcg acctcagcca tcccttcctg attttccgct ttccagcgtt  13980
cggcacgcag acgacgggct tcattctgca tggttgtgct taccagaccg gagatattga  14040
catcatatat gccttgagca actgatagct gtcgctgtca actgtcactg taatacgctg  14100
cttcatagca cacctctttt tgacatactt cgggtagtgc cgatcaacgt ctcattttcg  14160
ccaaaagttg gcccagggct tcccggtatc aacagggaca ccaggattta tttattctgc  14220
gaagtgatct tccgtcacag gtatttattc ggcgcaaagt gcgtcgggtg atgctgccaa  14280
cttagtcgac tacaggtcac taataccatc taagtagttg attcatagtg actggatatg  14340
```

```
ttgtgtttta cagtattatg tagtctgttt tttatgcaaa atctaattta atatattgat  14400
atttatatca ttttacgttt ctcgttcagc tttcttgtac aaagtggtct cgaggaattc  14460
ggtaccccag cttggtaagg aaataattat tttctttttt ccttttagta taaaatagtt  14520
aagtgatgtt aattagtatg attataataa tatagttgtt ataattgtga aaaaataatt  14580
tataaatata ttgtttacat aaacaacata gtaatgtaaa aaaatatgac aagtgatgtg  14640
taagacgaag aagataaaag ttgagagtaa gtatattatt tttaatgaat ttgatcgaac  14700
atgtaagatg atatactagc attaatattt gttttaatca taatagtaat tctagctggt  14760
ttgatgaatt aaatatcaat gataaaatac tatagtaaaa ataagaataa ataaattaaa  14820
ataatatttt tttatgatta atagtttatt atataattaa atatctatac cattactaaa  14880
tattttagtt taaaagttaa taaatatttt gttagaaatt ccaatctgct tgtaatttat  14940
caataaacaa aatattaaat aacaagctaa agtaacaaat aatatcaaac taatagaaac  15000
agtaatctaa tgtaacaaaa cataatctaa tgctaatata acaaagcgca agatctatca  15060
ttttatatag tattattttc aatcaacatt cttattaatt tctaaataat acttgtagtt  15120
ttattaactt ctaaatggat tgactattaa ttaaatgaat tagtcgaaca tgaataaaca  15180
aggtaacatg atagatcatg tcattgtgtt atcattgatc ttacatttgg attgattaca  15240
gttgggaagc tgggttcgaa atcgataagc ttggatcctc tagaccactt tgtacaagaa  15300
agctgaacga gaaacgtaaa atgatataaa tatcaatata ttaaattaga ttttgcataa  15360
aaaacagact acataatact gtaaaacaca acatatccag tcactatgaa tcaactactt  15420
agatggtatt agtgacctgt agtcgactaa gttggcagca tcacccgacg cactttgcgc  15480
cgaataaata cctgtgacgg aagatcactt cgcagaataa ataaatcctg gtgtccctgt  15540
tgataccggg aagccctggg ccaacttttg gcgaaaatga gacgttgatc ggatttcaca  15600
actcttatac ttttctctta caagtcgttc ggcttcatct ggattttcag cctctatact  15660
tactaaacgt gataaagttt ctgtaatttc tactgtatcg acctgcagac tggctgtgta  15720
taagggagcc tgacatttat attccccaga acatcaggtt aatggcgttt ttgatgtcat  15780
tttcgcggtg gctgagatca gccacttctt ccccgataac ggagaccggc acactggcca  15840
tatcggtggt catcatgcgc cagctttcat ccccgatatg caccaccggg taaagttcac  15900
gggagacttt atctgacagc agacgtgcac tggccagggg gatcaccatc cgtcgcccgg  15960
gcgtgtcaat aatatcactc tgtacatcca caaacagacg ataacggctc tctcttttat  16020
aggtgtaaac cttaaactgc atttcaccag tccctgttct cgtcagcaaa agagccgttc  16080
atttcaataa accgggcgac ctcagccatc ccttcctgat tttccgcttt ccagcgttcg  16140
gcacgcagac gacgggcttc attctgcatg gttgtgctta ccagaccgga gatattgaca  16200
tcatatatgc cttgagcaac tgatagctgt cgctgtcaac tgtcactgta atacgctgct  16260
tcatagcaca cctctttttg acatacttct gttcttgatg cagatgattt tcaggactat  16320
gacactagcg tatatgaata ggtagatgtt tttattttgt cacacaaaaa agaggctcgc  16380
acctcttttt cttatttctt tttatgattt aatacggcat tgaggacaat agcgagtagg  16440
ctggatacga cgattccgtt tgagaagaac atttggaagg ctgtcggtcg actaagttgg  16500
cagcatcacc cgaagaacat ttggaaggct gtcggtcgac tacaggtcac taataccatc  16560
taagtagttg attcatagtg actggatatg ttgtgtttta cagtattatg tagtctgttt  16620
tttatgcaaa atctaattta atatattgat atttatatca ttttacgttt ctcgttcagc  16680
ttttttgtac aaacttgtct agagtcctgc tttaatgaga tatgcgagac gcctatgatc  16740
gcatgatatt tgctttcaat tctgttgtgc acgttgtaaa aaacctgagc atgtgtagct  16800
cagatcctta ccgccggttt cggttcattc taatgaatat atcacccgtt actatcgtat  16860
ttttatgaat aatattctcc gttcaattta ctgattgtac cctactactt atatgtacaa  16920
tattaaaatg aaaacaatat attgtgctga ataggtttat agcgacatct atgatagagc  16980
gccacaataa caaacaattg cgttttatta ttacaaatcc aattttaaaa aaagcggcag  17040
aaccggtcaa acctaaaaga ctgattacat aaatcttatt caaatttcaa aaggccccag  17100
gggctagtat ctacgacaca ccgagcggcg aactaataac gttcactgaa gggaactccg  17160
gttccccgcc ggcgcgcatg ggtgagattc cttgaagttg agtattggcc gtccgctcta  17220
ccgaaagtta cgggcaccat tcaacccggt ccagcacggc ggccgggtaa ccgacttgct  17280
gccccgagaa ttatgcagca ttttttttggt gtatgtgggc cccaaatgaa gtgcaggtca  17340
aaccttgaca gtgacgacaa atcgttgggc gggtccaggg cgaattttgc gacaacatgt  17400
cgaggctcag caggacctgc aggcatgcaa gctagcttac tagtgatgca tattctatag  17460
tgtcacctaa atctgc                                                  17476
```

<210> 25
<211> 17458
<212> DNA
<213> Artificial sequence

<220>
<223> acceptor vector pHELLSGATE11

<400> 25

ggccgcacta gtgatatccc gcggccatgg cggccgggag catgcgacgt cgggcccaat 60
tcgccctata gtgagtcgta ttacaattca ctggccgtcg ttttacaacg tcgtgactgg 120
gaaaaccctg gcgttaccca acttaatcgc cttgcagcac atccccctttcgccagctgg 180
cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag cctgaatggc 240
gaatggaaat tgtaaacgtt aatgggtttc tggagtttaa tgagctaagc acatacgtca 300
gaaaccatta ttgcgcgttc aaaagtcgcc taaggtcact atcagctagc aaatatttct 360
tgtcaaaaat gctccactga cgttccataa attcccctcg gtatccaatt agagtctcat 420
attcactctc aatccaaata atctgcaatg gcaattacct tatccgcaac ttctttacct 480
atttccgccc ggatccgggc aggttctccg gccgcttggg tggagaggct attcggctat 540
gactgggcac aacagacaat cggctgctct gatgccgccg tgttccggct gtcagcgcag 600
gggcgcccgg ttctttttgt caagaccgac ctgtccggtg ccctgaatga actgcaggac 660
gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac 720
gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg gcaggatctc 780
ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc aatgcggcgg 840
ctgcatacgc ttgatccggc tacctgccca ttcgaccacc aagcgaaaca tcgcatcgag 900
cgagcacgta ctcggatgga agccggtctt gtcgatcagg atgatctgga cgaagagcat 960
caggggctcg cgccagccga actgttcgcc aggctcaagg cgcgcatgcc cgacggcgag 1020
gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc 1080
ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg 1140
ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg 1200
ctttacggta tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct tcttgacgag 1260
ttcttctgag cgggactctg gggttcgaaa tgaccgacca agcgacgccc aacctgccat 1320
cacgagattt cgattccacc gccgccttct atgaaaggtt gggcttcgga atcgttttcc 1380
gggacgccgg ctggatgatc ctccagcgcg gggatctcat gctggagttc ttcgcccacc 1440
ccgatccaac acttacgttt gcaacgtcca agagcaaata gaccacgaac gccggaaggt 1500
tgccgcagcg tgtggattgc gtctcaattc tctcttgcag gaatgcaatg atgaatatga 1560
tactgactat gaaactttga gggaatactg cctagcaccg tcacctcata acgtgcatca 1620
tgcatgccct gacaacatgg aacatcgcta tttttctgaa gaattatgct cgttggagga 1680
tgtcgcggca attgcagcta ttgccaacat cgaactaccc ctcacgcatg cattcatcaa 1740
tattattcat gcggggaaag gcaagattaa tccaactggc aaatcatcca gcgtgattgg 1800
taacttcagt tccagcgact tgattcgttt tggtgctacc cacgttttca ataaggacga 1860
gatggtggag taaagaagga gtgcgtcgaa gcagatcgtt caaacatttg gcaataaagt 1920
ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat 1980
tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag atgggttttt 2040
atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 2100
aactaggata aattatcgcg cgcggtgtca tctatgttac tagatcgaat taattccagg 2160
cggtgaaggg caatcagctg ttgcccgtct cactggtgaa aagaaaaacc accccagtac 2220
attaaaaacg tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa 2280
tatatcctgc caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca 2340
ctcgatacag gcagcccatc agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg 2400
cagactttgc tcatgttacc gatgctattc ggaagaacgg caactaagct gccgggtttg 2460
aaacacggat gatctcgcgg agggtagcat gttgattgta acgatgacag agcgttgctg 2520
cctgtgatca aatatcatct ccctcgcaga gatccgaatt atcagccttc ttattcattt 2580
ctcgcttaac cgtgacaggc tgtcgatctt gagaactatg ccgacataat aggaaatcgc 2640
tggataaagc cgctgaggaa gctgagtggc gctatttctt tagaagtgaa cgttgacgat 2700
gtcgacggat cttttccgct gcataaccct gcttcggggt cattatagcg atttttcgg 2760
tatatccatc ctttttcgca cgatatacag gattttgcca aagggttcgt gtagactttc 2820
cttggtgtat ccaacggcgt cagccgggca ggataggtga agtaggccca cccgcgagcg 2880
ggtgttcctt cttcactgtc ccttattcgc acctggcggt gctcaacggg aatcctgctc 2940
tgcgaggctg gccggctacc gccggcgtaa cagatgaggg caagcggatg gctgatgaaa 3000
ccaagccaac caggggtgat gctgccaact tactgattta gtgtatgatg gtgtttttga 3060
ggtgctccag tggcttctgt ttctatcagc tgtccctcct gttcagctac tgacggggtg 3120
gtgcgtaacg gcaaaagcac cgccggacat cagcgctatc tctgctctca ctgccgtaaa 3180
acatggcaac tgcagttcac ttacaccgct tctcaacccg gtacgcacca gaaaatcatt 3240
gatatggcca tgaatggcgt tggatgccgg gcaacagccc gcattatggg cgttggcctc 3300
aacacgattt tacgtcactt aaaaaactca ggccgcagtc ggtaacctcg cgcatacagc 3360
cgggcagtga cgtcatcgtc tgcgcggaaa tggacgaaca gtggggctat gtcggggcta 3420
aatcgcgcca gcgctggctg ttttacgcgt atgacagtct ccggaagacg gttgttgcgc 3480
acgtattcgg tgaacgcact atggcgacgc tggggcgtct tatgagcctg ctgtcaccct 3540
ttgacgtggt gatatggatg acggatggct ggccgctgta tgaatcccgc ctgaagggaa 3600
agctgcacgt aatcagcaag cgatatacgc agcgaattga gcggcataac ctgaatctga 3660
ggcagcacct ggcacggctg ggacggaagt cgctgtcgtt ctcaaaatcg gtggagctgc 3720

```
atgacaaagt catcgggcat tatctgaaca taaaacacta tcaataagtt ggagtcatta   3780
cccaaccagg aagggcagcc cacctatcaa ggtgtactgc cttccagacg aacgaagagc   3840
gattgaggaa aaggcggcgg cggccggcat gagcctgtcg gcctacctgc tggccgtcgg   3900
ccagggctac aaaatcacgg gcgtcgtgga ctatgagcac gtccgcgagc tggcccgcat   3960
caatggcgac ctgggccgcc tgggcggcct gctgaaactc tggctcaccg acgacccgcg   4020
cacggcgcgg ttcggtgatg ccacgatcct cgccctgctg gcgaagatcg aagagaagca   4080
ggacgagctt ggcaaggtca tgatgggcgt ggtccgcccg agggcagagc catgacttttt   4140
ttagccgcta aaacggccgg ggggtgcgcg tgattgccaa gcacgtcccc atgcgctcca   4200
tcaagaagag cgacttcgcg gagctggtat tcgtgcaggg caagattcgg aataccaagt   4260
acgagaagga cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc   4320
tggacaccaa ggcaccaggc gggtcaaatc aggaataagg gcacattgcc ccggcgtgag   4380
tcggggcaat cccgcaagga gggtgaatga atcggacgtt tgaccggaag gcatacaggc   4440
aagaactgat cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg   4500
tcatgcgtgc gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg   4560
ccaagatcga gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg   4620
ccgtggagcg ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag tcgatgacca   4680
tcgacacgcg aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa   4740
aacaggtcag cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg   4800
aaatgcagct ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa   4860
acgacacggc ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc   4920
tgcaaaacaa ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg   4980
agctgcgggc cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca   5040
cccctatcgg cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt   5100
cgatcaatgg ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg   5160
cgatgggctt cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct   5220
tccgcgtcct ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa   5280
tcgtcgtgct gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc   5340
tgtcgccgac ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc   5400
tcaagctgga aaccttccgc ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc   5460
gcgagcaggt cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct   5520
gggtcaatga tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg   5580
ggggttcagc agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca   5640
cttgcttcgc tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag   5700
aggattaaaa ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt   5760
gcaggatttc cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc   5820
cgtttacgag cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc   5880
cgtggcattc ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga   5940
ggacggcccc aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca   6000
gcgaggccga ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt   6060
gatgatcgtc cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc   6120
acttaatatt tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg   6180
ggtcgcggcg acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct   6240
gctaggtagc ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt   6300
ggcgctgttg gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct   6360
ggcgggggcg gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt   6420
gcctctgctc acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt   6480
agctttagtg tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc   6540
gtggctcggc ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg   6600
actcgaacct acagttgttt ccttactggg ctttctcagc cgggatggcg ctaagaagct   6660
attgccgccg atcttcatat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac   6720
cgcatcaggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg   6780
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat   6840
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc   6900
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacaa aaatcgacgc   6960
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga   7020
agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt   7080
ctcccttcgg gaagcgtggc gctttctcaa tgctcacgct gtaggtatct cagttcggtg   7140
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc   7200
gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg   7260
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc   7320
ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg   7380
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc   7440
gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatat   7500
```

caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt 7560
taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa 7620
aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa 7680
tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc 7740
tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct 7800
gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca 7860
gccggaaggg ccgagcgcag aagtggtcct gcaactttat ccgcctccat ccagtctatt 7920
aaacaagtgg cagcaacgga ttcgcaaacc tgtcacgcct tttgtgccaa aagccgcgcc 7980
aggtttgcga tccgctgtgc caggcgttag gcgtcatatg aagatttcgg tgatccctga 8040
gcaggtggcg gaaacattgg atgctgagaa ccatttcatt gttcgtgaag tgttcgatgt 8100
gcacctatcc gaccaaggct ttgaactatc taccagaagt gtgagcccct accggaagga 8160
ttacatctcg gatgatgact ctgatgaaga ctctgcttgc tatggcgcat tcatcgacca 8220
agagcttgtc gggaagattg aactcaactc aacatggaac gatctagcct ctatcgaaca 8280
cattgttgtg tcgcacacgc accgaggcaa aggagtcgcg cacagtctca tcgaatttgc 8340
gaaaaagtgg gcactaagca gacagctcct tggcatacga ttagagacac aaacgaacaa 8400
tgtacctgcc tgcaatttgt acgcaaaatg tggctttact ctcggcggca ttgacctgtt 8460
cacgtataaa actagacctc aagtctcgaa cgaaacagcg atgtactggt actggttctc 8520
gggagcacag gatgacgcct aacaattcat tcaagccgac accgcttcgc ggcgcggctt 8580
aattcaggag ttaaacatca tgagggaagc ggtgatcgcc gaagtatcga ctcaactatc 8640
agaggtagtt ggcgtcatcg agcgccatct cgaaccgacg ttgctggccg tacatttgta 8700
cggctccgca gtggatggcg gcctgaagcc acacagtgat attgatttgc tggttacggt 8760
gaccgtaagg cttgatgaaa caacgcggcg agctttgatc aacgaccttt tggaaacttc 8820
ggcttcccct ggagagagcg agattctccg cgctgtagaa gtcaccattg ttgtgcacga 8880
cgacatcatt ccgtggcgtt atccagctaa gcgcgaactg caatttggag aatggcagcg 8940
caatgacatt cttgcaggta tcttcgagcc agccacgatc gacattgatc tggctatctt 9000
gctgacaaaa gcaagagaac atagcgttgc cttggtaggt ccagcggcgg aggaactctt 9060
tgatccggtt cctgaacagg atctatttga ggcgctaaat gaaaccttaa cgctatggaa 9120
ctcgccgccc gactgggctg gcgatgagcg aaatgtagtg cttacgttgt cccgcatttg 9180
gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc gctgccgact gggcaatgga 9240
gcgcctgccg gcccagtatc agcccgtcat acttgaagct aggcaggctt atcttggaca 9300
agaagatcgc ttggcctcgc gcgcagatca gttggaagaa tttgttcact acgtgaaagg 9360
cgagatcacc aaggtagtcg gcaaataatg tctaacaatt cgttcaagcc gacgccgctt 9420
cgcggcgcgg cttaactcaa gcgttagaga gctggggaag actatgcgcg atctgttgaa 9480
ggtggttcta agcctcgtac ttgcgatggc atcggggcag gcacttgctg acctgccaat 9540
tgttttagtg gatgaagctc gtcttcccta tgactactcc ccatccaact acgacatttc 9600
tccaagcaac tacgacaact ccataagcaa ttacgacaat agtccatcaa attacgacaa 9660
ctctgagagc aactacgata atagttcatc caattacgac aatagtcgca acggaaatcg 9720
taggcttata tatagcgcaa atgggtctcg cactttcgcc ggctactacg tcattgccaa 9780
caatgggaca acgaacttct tttccacatc tggcaaaagg atgttctaca ccccaaaagg 9840
ggggcgcggc gtctatggcg gcaaagatgg gagcttctgc ggggcattgg tcgtcataaa 9900
tggccaattt tcgcttgccc tgacagataa cggcctgaag atcatgtatc taagcaacta 9960
gcctgctctc taataaaatg ttaggagctt ggctgccatt tttggggtga ggccgttcgc 10020
ggccgagggg cgcagcccct ggggggatgg gaggcccgcg ttagcgggcc gggagggttc 10080
gagaaggggg ggcacccccc ttcggcgtgc gcggtcacgc gccagggcgc agccctggtt 10140
aaaaacaagg tttataaata ttggtttaaa agcaggttaa aagacaggtt agcggtggcc 10200
gaaaaacggg cggaaaccct tgcaaatgct ggattttctg cctgtggaca gcccctcaaa 10260
tgtcaatagg tgcgcccctc atctgtcagc actctgcccc tcaagtgtca aggatcgcgc 10320
ccctcatctg tcagtagtcg cgcccctcaa gtgtcaatac cgcagggcac ttatccccag 10380
gcttgtccac atcatctgtg ggaaactcgc gtaaaatcag gcgttttcgc cgatttgcga 10440
ggctggccag ctccacgtcg ccggccgaaa tcgagcctgc ccctcatctg tcaacgccgc 10500
gccgggtgag tcggcccctc aagtgtcaac gtccgccccc catctgtcag tgagggccaa 10560
gttttccgcg aggtatccac aacgccggcg gccggccgcg gtgtctcgca cacggcttcg 10620
acggcgtttc tggcgcgttt gcagggccat agacggccgc cagcccagcg gcgagggcaa 10680
ccagcccggt gagcgtcgga aagggtcgac atcttgctgc gttcggatat tttcgtggag 10740
ttcccgccac agacccggat tgaaggcgag atccagcaac tcgcgccaga tcatcctgtg 10800
acggaacttt ggcgcgtgat gactggccag gacgtcggcc gaaagagcga caagcagatc 10860
acgattttcg acagcgtcgg atttgcgatc gaggattttt cggcgctgcg ctacgtccgc 10920
gaccgcgttg agggatcaag ccacagcagc ccactcgacc ttctagccga cccagacgag 10980
ccaagggatc tttttggaat gctgctccgt cgtcaggctt tccgacgttt gggtggttga 11040
acagaagtca ttatcgtacg gaatgccagc actcccgagg ggaaccctgt ggttggcatg 11100
cacatacaaa tggacgaacg gataaacctt ttcacgccct tttaaatatc cgttattcta 11160
ataaacgctc ttttctctta ggtttacccg ccaatatatc ctgtcaaaca ctgatagttt 11220
aaactgaagg cgggaaacga caatctgatc atgagcggag aattaaggga gtcacgttat 11280

gacccccgcc gatgacgcgg gacaagccgt tttacgtttg gaactgacag aaccgcaacg 11340
attgaaggag ccactcagcc ccaatacgca aaccgcctct ccccgcgcgt tggccgattc 11400
attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa 11460
ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg cttccggctc 11520
gtatgttgtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg 11580
attacgccaa gctatttagg tgacactata gaatactcaa gctatgcatc caacgcgttg 11640
ggagctctcc catatcgacc tgcaggcggc cgctcgacga attaattcca atcccacaaa 11700
aatctgagct taacagcaca gttgctcctc tcagagcaga atcgggtatt caacaccctc 11760
atatcaacta ctacgttgtg tataacggtc cacatgccgg tatatacgat gactggggtt 11820
gtacaaaggc ggcaacaaac ggcgttcccg gagttgcaca caagaaattt gccactatta 11880
cagaggcaag agcagcagct gacgcgtaca caacaagtca gcaaacagac aggttgaact 11940
tcatccccaa aggagaagct caactcaagc ccaagagctt tgctaaggcc ctaacaagcc 12000
caccaaagca aaaagcccac tggctcacgc taggaaccaa aaggcccagc agtgatccag 12060
ccccaaaaga gatctccttt gccccggaga ttacaatgga cgatttcctc tatctttacg 12120
atctaggaag gaagttcgaa ggtgaaggtg acgacactat gttcaccact gataatgaga 12180
aggttagcct cttcaatttc agaaagaatg ctgacccaca gatggttaga gaggcctacg 12240
cagcaggtct catcaagacg atctacccga gtaacaatct ccaggagatc aaataccttc 12300
ccaagaaggt taaagatgca gtcaaaagat tcaggactaa ttgcatcaag aacacagaga 12360
aagacatatt tctcaagatc agaagtacta ttccagtatg gacgattcaa ggcttgcttc 12420
ataaaccaag gcaagtaata gagattggag tctctaaaaa ggtagttcct actgaatcta 12480
aggccatgca tggagtctaa gattcaaatc gaggatctaa cagaactcgc cgtgaagact 12540
ggcgaacagt tcatacagag tcttttacga ctcaatgaca agaagaaaat cttcgtcaac 12600
atggtggagc acgacactct ggtctactcc aaaaatgtca aagatacagt ctcagaagac 12660
caaagggcta ttgagacttt tcaacaaagg ataatttcgg gaaacctcct cggattccat 12720
tgcccagcta tctgtcactt catcgaaagg acagtagaaa aggaaggtgg ctcctacaaa 12780
tgccatcatt gcgataaagg aaaggctatc attcaagatc tctctgccga cagtggtccc 12840
aaagatggac ccccacccac gaggagcatc gtggaaaaag aagacgttcc aaccacgtct 12900
tcaaagcaag tggattgatg tgacatctcc actgacgtaa gggatgacgc acaatcccac 12960
tatccttcgc aagacccttc ctctatataa ggaagttcat ttcatttgga gaggacacgc 13020
tcgagacaag tttgtacaaa aaagctgaac gagaaacgta aaatgatata aatatcaata 13080
tattaaatta gattttgcat aaaaaacaga ctacataata ctgtaaaaca caacatatcc 13140
agtcactatg aatcaactac ttagatggta ttagtgacct gtagtcgacc gacagccttc 13200
caaatgttct tcgggtgatg ctgccaactt agtcgaccga cagccttcca aatgttcttc 13260
tcaaacggaa tcgtcgtatc cagcctactc gctattgtcc tcaatgccgt attaaatcat 13320
aaaaagaaat aagaaaaaga ggtgcgagcc tcttttttgt gtgacaaaat aaaaacatct 13380
acctattcat atacgctagt gtcatagtcc tgaaaatcat ctgcatcaag aacaatttca 13440
caactcttat acttttctct tacaagtcgt tcggcttcat ctggattttc agcctctata 13500
cttactaaac gtgataaagt ttctgtaatt tctactgtat cgacctgcag actggctgtg 13560
tataagggag cctgacattt atattcccca gaacatcagg ttaatggcgt ttttgatgtc 13620
attttcgcgg tggctgagat cagccacttc ttccccgata acggagaccg gcacactggc 13680
catatcggtg gtcatcatgc gccagctttc atccccgata tgcaccaccg ggtaaagttc 13740
acgggagact ttatctgaca gcagacgtgc actggccagg gggatcacca tccgtcgccc 13800
gggcgtgtca ataatatcac tctgtacatc cacaaacaga cgataacggc tctctctttt 13860
ataggtgtaa accttaaact gcatttcacc agtccctgtt ctcgtcagca aaagagccgt 13920
tcatttcaat aaaccgggcg acctcagcca tcccttcctg attttccgct ttccagcgtt 13980
cggcacgcag acgacgggct tcattctgca tggttgtgct taccagaccg gagatattga 14040
catcatatat gccttgagca actgatagct gtcgctgtca actgtcactg taatacgctg 14100
cttcatagca cacctctttt tgacatactt cgggtagtgc cgatcaacgt ctcattttcg 14160
ccaaaagttg gcccagggct tcccggtatc aacagggaca ccaggattta tttattctgc 14220
gaagtgatct tccgtcacag gtatttattc ggcgcaaagt gcgtcgggtg atgctgccaa 14280
cttagtcgac tacaggtcac taataccatc taagtagttg attcatagtg actggatatg 14340
ttgtgtttta cagtattatg tagtctgttt tttatgcaaa atctaattta atatattgat 14400
atttatatca ttttacgttt ctcgttcagc tttcttgtac aaagtggtct cgaggaattc 14460
ggtaccaact gtaaggaaat aattattttc tttttccctt ttagtataaa atagttaagt 14520
gatgttaatt agtatgatta taataatata gttgttataa ttgtgaaaaa ataatttata 14580
aatatattgt ttacataaac aacatagtaa tgtaaaaaaa tatgacaagt gatgtgtaag 14640
acgaagaaga taaaagttga gagtaagtat attattttta atgaatttga tcgaacatgt 14700
aagatgatat actagcatta atatttgttt taatcataat agtaattcta gctggtttga 14760
tgaattaaat atcaatgata aaatactata gtaaaaataa gaataaataa attaaaataa 14820
tattttttta tgattaatag tttattatat aattaaatat ctataccatt actaaatatt 14880
ttagtttaaa agttaataaa tattttgtta gaaattccaa tctgcttgta atttatcaat 14940
aaacaaaata ttaaataaca agctaaagta acaaataata tcaaactaat agaaacagta 15000
atctaatgta acaaaacata atctaatgct aatataacaa agcgcaagat ctatcatttt 15060

atatagtatt attttcaatc aacattctta ttaatttcta aataatactt gtagttttat 15120
taacttctaa atggattgac tattaattaa atgaattagt cgaacatgaa taaacaaggt 15180
aacatgatag atcatgtcat tgtgttatca ttgatcttac atttggattg attacagtta 15240
cttaccttaa gcttggatcc tctagaccac tttgtacaag aaagctgaac gagaaacgta 15300
aaatgatata aatatcaata tattaaatta gattttgcat aaaaaacaga ctacataata 15360
ctgtaaaaca caacatatcc agtcactatg aatcaactac ttagatggta ttagtgacct 15420
gtagtcgact aagttggcag catcacccga cgcactttgc gccgaataaa tacctgtgac 15480
ggaagatcac ttcgcagaat aaataaatcc tggtgtccct gttgataccg ggaagccctg 15540
ggccaacttt tggcgaaaat gagacgttga tcggatttca caactcttat acttttctct 15600
tacaagtcgt tcggcttcat ctggattttc agcctctata cttactaaac gtgataaagt 15660
ttctgtaatt tctactgtat cgacctgcag actggctgtg tataagggag cctgacattt 15720
atattcccca gaacatcagg ttaatggcgt ttttgatgtc attttcgcgg tggctgagat 15780
cagccacttc ttccccgata acggagaccg gcacactggc catatcggtg gtcatcatgc 15840
gccagctttc atccccgata tgcaccaccg ggtaaagttc acgggagact ttatctgaca 15900
gcagacgtgc actggccagg gggatcacca tccgtcgccc gggcgtgtca ataatatcac 15960
tctgtacatc cacaaacaga cgataacggc tctctctttt ataggtgtaa accttaaact 16020
gcatttcacc agtccctgtt ctcgtcagca aaagagccgt tcatttcaat aaaccgggcg 16080
acctcagcca tcccttcctg attttccgct ttccagcgtt cggcacgcag acgacgggct 16140
tcattctgca tggttgtgct taccagaccg gagatattga catcatatat gccttgagca 16200
actgatagct gtcgctgtca actgtcactg taatacgctg cttcatagca cacctctttt 16260
tgacatactt ctgttcttga tgcagatgat tttcaggact atgacactag cgtatatgaa 16320
taggtagatg tttttatttt gtcacacaaa aaagaggctc gcacctcttt ttcttatttc 16380
tttttatgat ttaatacggc attgaggaca atagcgagta ggctggatac gacgattccg 16440
tttgagaaga acatttggaa ggctgtcggt cgactaagtt ggcagcatca cccgaagaac 16500
atttggaagg ctgtcggtcg actacaggtc actaatacca tctaagtagt tgattcatag 16560
tgactggata tgttgtgttt tacagtatta tgtagtctgt tttttatgca aaatctaatt 16620
taatatattg atatttatat cattttacgt ttctcgttca gctttttgt acaaacttgt 16680
ctagagtcct gctttaatga gatatgcgag acgcctatga tcgcatgata tttgctttca 16740
attctgttgt gcacgttgta aaaaacctga gcatgtgtag ctcagatcct taccgccggt 16800
ttcggttcat tctaatgaat atatcacccg ttactatcgt atttttatga ataatattct 16860
ccgttcaatt tactgattgt accctactac ttatatgtac aatattaaaa tgaaaacaat 16920
atattgtgct gaataggttt atagcgacat ctatgataga gcgccacaat aacaaacaat 16980
tgcgttttat tattacaaat ccaattttaa aaaaagcggc agaaccggtc aaacctaaaa 17040
gactgattac ataaatctta ttcaaatttc aaaaggcccc aggggctagt atctacgaca 17100
caccgagcgg cgaactaata acgttcactg aagggaactc cggttccccg ccggcgcgca 17160
tgggtgagat tccttgaagt tgagtattgg ccgtccgctc taccgaaagt tacgggcacc 17220
attcaacccg gtccagcacg gcggccgggt aaccgacttg ctgccccgag aattatgcag 17280
catttttttg gtgtatgtgg gccccaaatg aagtgcaggt caaaccttga cagtgacgac 17340
aaatcgttgg gcgggtccag ggcgaatttt gcgacaacat gtcgaggctc agcaggacct 17400
gcaggcatgc aagctagctt actagtgatg catattctat agtgtcacct aaatctgc 17458

<210> 26
<211> 17681
<212> DNA
<213> Artificial sequence

<220>
<223> acceptor vector pHELLSGATE12

<400> 26
ggccgcacta gtgatatccc gcggccatgg cggccgggag catgcgacgt cgggcccaat 60
tcgccctata gtgagtcgta ttacaattca ctggccgtcg ttttacaacg tcgtgactgg 120
gaaaaccctg gcgttaccca acttaatcgc cttgcagcac atccccctttcgccagctgg 180
cgtaatagcg aagaggcccg caccgatcgc ccttcccaac agttgcgcag cctgaatggc 240
gaatggaaat tgtaaacgtt aatgggtttc tggagtttaa tgagctaagc acatacgtca 300
gaaaccatta ttgcgcgttc aaaagtcgcc taaggtcact atcagctagc aaatatttct 360
tgtcaaaaat gctccactga cgttccataa attcccctcg gtatccaatt agagtctcat 420
attcactctc aatccaaata atctgcaatg gcaattacct tatccgcaac ttctttacct 480
atttccgccc ggatccgggc aggttctccg gccgcttggg tggagaggct attcggctat 540
gactgggcac aacagacaat cggctgctct gatgccgccg tgttccggct gtcagcgcag 600
gggcgcccgg ttctttttgt caagaccgac ctgtccggtg ccctgaatga actgcaggac 660
gaggcagcgc ggctatcgtg gctggccacg acgggcgttc cttgcgcagc tgtgctcgac 720

gttgtcactg aagcgggaag ggactggctg ctattgggcg aagtgccggg gcaggatctc 780
ctgtcatctc accttgctcc tgccgagaaa gtatccatca tggctgatgc aatgcggcgg 840
ctgcatacgc ttgatccggc tacctgccca ttcgaccacc aagcgaaaca tcgcatcgag 900
cgagcacgta ctcggatgga agccggtctt gtcgatcagg atgatctgga cgaagagcat 960
caggggctcg cgccagccga actgttcgcc aggctcaagg cgcgcatgcc cgacggcgag 1020
gatctcgtcg tgacccatgg cgatgcctgc ttgccgaata tcatggtgga aaatggccgc 1080
ttttctggat tcatcgactg tggccggctg ggtgtggcgg accgctatca ggacatagcg 1140
ttggctaccc gtgatattgc tgaagagctt ggcggcgaat gggctgaccg cttcctcgtg 1200
ctttacggta tcgccgctcc cgattcgcag cgcatcgcct tctatcgcct tcttgacgag 1260
ttcttctgag cgggactctg gggttcgaaa tgaccgacca agcgacgccc aacctgccat 1320
cacgagattt cgattccacc gccgccttct atgaaaggtt gggcttcgga atcgttttcc 1380
gggacgccgg ctggatgatc ctccagcgcg gggatctcat gctggagttc ttcgcccacc 1440
ccgatccaac acttacgttt gcaacgtcca agagcaaata gaccacgaac gccggaaggt 1500
tgccgcagcg tgtggattgc gtctcaattc tctcttgcag gaatgcaatg atgaatatga 1560
tactgactat gaaactttga gggaatactg cctagcaccg tcacctcata acgtgcatca 1620
tgcatgccct gacaacatgg aacatcgcta tttttctgaa gaattatgct cgttggagga 1680
tgtcgcggca attgcagcta ttgccaacat cgaactaccc ctcacgcatg cattcatcaa 1740
tattattcat gcggggaaag gcaagattaa tccaactggc aaatcatcca gcgtgattgg 1800
taacttcagt tccagcgact tgattcgttt tggtgctacc cacgttttca ataaggacga 1860
gatggtggag taaagaagga gtgcgtcgaa gcagatcgtt caaacatttg gcaataaagt 1920
ttcttaagat tgaatcctgt tgccggtctt gcgatgatta tcatataatt tctgttgaat 1980
tacgttaagc atgtaataat taacatgtaa tgcatgacgt tatttatgag atgggttttt 2040
atgattagag tcccgcaatt atacatttaa tacgcgatag aaaacaaaat atagcgcgca 2100
aactaggata aattatcgcg cgcggtgtca tctatgttac tagatcgaat taattccagg 2160
cggtgaaggg caatcagctg ttgcccgtct cactggtgaa aagaaaaacc accccagtac 2220
attaaaaacg tccgcaatgt gttattaagt tgtctaagcg tcaatttgtt tacaccacaa 2280
tatatcctgc caccagccag ccaacagctc cccgaccggc agctcggcac aaaatcacca 2340
ctcgatacag gcagcccatc agtccgggac ggcgtcagcg ggagagccgt tgtaaggcgg 2400
cagactttgc tcatgttacc gatgctattc ggaagaacgg caactaagct gccgggtttg 2460
aaacacggat gatctcgcgg agggtagcat gttgattgta acgatgacag agcgttgctg 2520
cctgtgatca aatatcatct ccctcgcaga gatccgaatt atcagccttc ttattcattt 2580
ctcgcttaac cgtgacaggc tgtcgatctt gagaactatg ccgacataat aggaaatcgc 2640
tggataaagc cgctgaggaa gctgagtggc gctatttctt tagaagtgaa cgttgacgat 2700
gtcgacggat cttttccgct gcataaccct gcttcggggt cattatagcg atttttcgg 2760
tatatccatc ctttttcgca cgatatacag gattttgcca aagggttcgt gtagactttc 2820
cttggtgtat ccaacggcgt cagccgggca ggataggtga agtaggccca cccgcgagcg 2880
ggtgttcctt cttcactgtc ccttattcgc acctggcggt gctcaacggg aatcctgctc 2940
tgcgaggctg gccggctacc gccggcgtaa cagatgaggg caagcggatg gctgatgaaa 3000
ccaagccaac caggggtgat gctgccaact tactgattta gtgtatgatg gtgtttttga 3060
ggtgctccag tggcttctgt ttctatcagc tgtccctcct gttcagctac tgacggggtg 3120
gtgcgtaacg gcaaaagcac cgccggacat cagcgctatc tctgctctca ctgccgtaaa 3180
acatggcaac tgcagttcac ttacaccgct tctcaacccg gtacgcacca gaaaatcatt 3240
gatatggcca tgaatggcgt tggatgccgg gcaacagccc gcattatggg cgttggcctc 3300
aacacgattt tacgtcactt aaaaaactca ggccgcagtc ggtaacctcg cgcatacagc 3360
cgggcagtga cgtcatcgtc tgcgcggaaa tggacgaaca gtggggctat gtcggggcta 3420
aatcgcgcca gcgctggctg ttttacgcgt atgacagtct ccggaagacg gttgttgcgc 3480
acgtattcgg tgaacgcact atggcgacgc tggggcgtct tatgagcctg ctgtcaccct 3540
ttgacgtggt gatatggatg acggatggct ggccgctgta tgaatcccgc ctgaagggaa 3600
agctgcacgt aatcagcaag cgatatacgc agcgaattga gcggcataac ctgaatctga 3660
ggcagcacct ggcacggctg ggacggaagt cgctgtcgtt ctcaaaatcg gtggagctgc 3720
atgacaaagt catcgggcat tatctgaaca taaaacacta tcaataagtt ggagtcatta 3780
cccaaccagg aagggcagcc cacctatcaa ggtgtactgc cttccagacg aacgaagagc 3840
gattgaggaa aaggcggcgg cggccggcat gagcctgtcg gcctacctgc tggccgtcgg 3900
ccagggctac aaaatcacgg gcgtcgtgga ctatgagcac gtccgcgagc tggcccgcat 3960
caatggcgac ctgggccgcc tgggcggcct gctgaaactc tggctcaccg acgacccgcg 4020
cacggcgcgg ttcggtgatg ccacgatcct cgccctgctg gcgaagatcg aagagaagca 4080
ggacgagctt ggcaaggtca tgatgggcgt ggtccgcccg agggcagagc catgactttt 4140
ttagccgcta aaacggccgg ggggtgcgcg tgattgccaa gcacgtcccc atgcgctcca 4200
tcaagaagag cgacttcgcg gagctggtat tcgtgcaggg caagattcgg aataccaagt 4260
acgagaagga cggccagacg gtctacggga ccgacttcat tgccgataag gtggattatc 4320
tggacaccaa ggcaccaggc gggtcaaatc aggaataagg gcacattgcc ccggcgtgag 4380
tcggggcaat cccgcaagga gggtgaatga atcggacgtt tgaccggaag gcatacaggc 4440
aagaactgat cgacgcgggg ttttccgccg aggatgccga aaccatcgca agccgcaccg 4500

```
tcatgcgtgc gccccgcgaa accttccagt ccgtcggctc gatggtccag caagctacgg   4560
ccaagatcga gcgcgacagc gtgcaactgg ctccccctgc cctgcccgcg ccatcggccg   4620
ccgtggagcg ttcgcgtcgt ctcgaacagg aggcggcagg tttggcgaag tcgatgacca   4680
tcgacacgcg aggaactatg acgaccaaga agcgaaaaac cgccggcgag gacctggcaa   4740
aacaggtcag cgaggccaag caggccgcgt tgctgaaaca cacgaagcag cagatcaagg   4800
aaatgcagct ttccttgttc gatattgcgc cgtggccgga cacgatgcga gcgatgccaa   4860
acgacacggc ccgctctgcc ctgttcacca cgcgcaacaa gaaaatcccg cgcgaggcgc   4920
tgcaaaacaa ggtcattttc cacgtcaaca aggacgtgaa gatcacctac accggcgtcg   4980
agctgcgggc cgacgatgac gaactggtgt ggcagcaggt gttggagtac gcgaagcgca   5040
cccctatcgg cgagccgatc accttcacgt tctacgagct ttgccaggac ctgggctggt   5100
cgatcaatgg ccggtattac acgaaggccg aggaatgcct gtcgcgccta caggcgacgg   5160
cgatgggctt cacgtccgac cgcgttgggc acctggaatc ggtgtcgctg ctgcaccgct   5220
tccgcgtcct ggaccgtggc aagaaaacgt cccgttgcca ggtcctgatc gacgaggaaa   5280
tcgtcgtgct gtttgctggc gaccactaca cgaaattcat atgggagaag taccgcaagc   5340
tgtcgccgac ggcccgacgg atgttcgact atttcagctc gcaccgggag ccgtacccgc   5400
tcaagctgga aaccttccgc ctcatgtgcg gatcggattc cacccgcgtg aagaagtggc   5460
gcgagcaggt cggcgaagcc tgcgaagagt tgcgaggcag cggcctggtg gaacacgcct   5520
gggtcaatga tgacctggtg cattgcaaac gctagggcct tgtggggtca gttccggctg   5580
ggggttcagc agccagcgct ttactggcat ttcaggaaca agcgggcact gctcgacgca   5640
cttgcttcgc tcagtatcgc tcgggacgca cggcgcgctc tacgaactgc cgataaacag   5700
aggattaaaa ttgacaattg tgattaaggc tcagattcga cggcttggag cggccgacgt   5760
gcaggatttc cgcgagatcc gattgtcggc cctgaagaaa gctccagaga tgttcgggtc   5820
cgtttacgag cacgaggaga aaaagcccat ggaggcgttc gctgaacggt tgcgagatgc   5880
cgtggcattc ggcgcctaca tcgacggcga gatcattggg ctgtcggtct tcaaacagga   5940
ggacggcccc aaggacgctc acaaggcgca tctgtccggc gttttcgtgg agcccgaaca   6000
gcgaggccga ggggtcgccg gtatgctgct gcgggcgttg ccggcgggtt tattgctcgt   6060
gatgatcgtc cgacagattc caacgggaat ctggtggatg cgcatcttca tcctcggcgc   6120
acttaatatt tcgctattct ggagcttgtt gtttatttcg gtctaccgcc tgccgggcgg   6180
ggtcgcggcg acggtaggcg ctgtgcagcc gctgatggtc gtgttcatct ctgccgctct   6240
gctaggtagc ccgatacgat tgatggcggt cctgggggct atttgcggaa ctgcgggcgt   6300
ggcgctgttg gtgttgacac caaacgcagc gctagatcct gtcggcgtcg cagcgggcct   6360
ggcgggggcg gtttccatgg cgttcggaac cgtgctgacc cgcaagtggc aacctcccgt   6420
gcctctgctc acctttaccg cctggcaact ggcggccgga ggacttctgc tcgttccagt   6480
agctttagtg tttgatccgc caatcccgat gcctacagga accaatgttc tcggcctggc   6540
gtggctcggc ctgatcggag cgggtttaac ctacttcctt tggttccggg ggatctcgcg   6600
actcgaacct acagttgttt ccttactggg ctttctcagc cgggatggcg ctaagaagct   6660
attgccgccg atcttcatat gcggtgtgaa ataccgcaca gatgcgtaag gagaaaatac   6720
cgcatcaggc gctcttccgc ttcctcgctc actgactcgc tgcgctcggt cgttcggctg   6780
cggcgagcgg tatcagctca ctcaaaggcg gtaatacggt tatccacaga atcaggggat   6840
aacgcaggaa agaacatgtg agcaaaaggc cagcaaaagg ccaggaaccg taaaaaggcc   6900
gcgttgctgg cgtttttcca taggctccgc cccctgacg agcatcacaa aaatcgacgc   6960
tcaagtcaga ggtggcgaaa cccgacagga ctataaagat accaggcgtt tccccctgga   7020
agctccctcg tgcgctctcc tgttccgacc ctgccgctta ccggatacct gtccgccttt   7080
ctcccttcgg gaagcgtggc gctttctcaa tgctcacgct gtaggtatct cagttcggtg   7140
taggtcgttc gctccaagct gggctgtgtg cacgaacccc ccgttcagcc cgaccgctgc   7200
gccttatccg gtaactatcg tcttgagtcc aacccggtaa gacacgactt atcgccactg   7260
gcagcagcca ctggtaacag gattagcaga gcgaggtatg taggcggtgc tacagagttc   7320
ttgaagtggt ggcctaacta cggctacact agaaggacag tatttggtat ctgcgctctg   7380
ctgaagccag ttaccttcgg aaaaagagtt ggtagctctt gatccggcaa acaaaccacc   7440
gctggtagcg gtggtttttt tgtttgcaag cagcagatta cgcgcagaaa aaaaggatat   7500
caagaagatc ctttgatctt ttctacgggg tctgacgctc agtggaacga aaactcacgt   7560
taagggattt tggtcatgag attatcaaaa aggatcttca cctagatcct tttaaattaa   7620
aaatgaagtt ttaaatcaat ctaaagtata tatgagtaaa cttggtctga cagttaccaa   7680
tgcttaatca gtgaggcacc tatctcagcg atctgtctat ttcgttcatc catagttgcc   7740
tgactccccg tcgtgtagat aactacgata cgggagggct taccatctgg ccccagtgct   7800
gcaatgatac cgcgagaccc acgctcaccg gctccagatt tatcagcaat aaaccagcca   7860
gccggaaggg ccgagcgcag aagtggtcct gcaacttttat ccgcctccat ccagtctatt   7920
aaacaagtgg cagcaacgga ttcgcaaacc tgtcacgcct tttgtgccaa aagccgcgcc   7980
aggtttgcga tccgctgtgc caggcgttag gcgtcatatg aagatttcgg tgatccctga   8040
gcaggtggcg gaaacattgg atgctgagaa ccatttcatt gttcgtgaag tgttcgatgt   8100
gcacctatcc gaccaaggct ttgaactatc taccagaagt gtgagccccct accggaagga   8160
ttacatctcg gatgatgact ctgatgaaga ctctgcttgc tatggcgcat tcatcgacca   8220
agagcttgtc gggaagattg aactcaactc aacatggaac gatctagcct ctatcgaaca   8280
```

cattgttgtg tcgcacacgc accgaggcaa aggagtcgcg cacagtctca tcgaatttgc 8340
gaaaaagtgg gcactaagca gacagctcct tggcatacga ttagagacac aaacgaacaa 8400
tgtacctgcc tgcaatttgt acgcaaaatg tggctttact ctcggcggca ttgacctgtt 8460
cacgtataaa actagacctc aagtctcgaa cgaaacagcg atgtactggt actggttctc 8520
gggagcacag gatgacgcct aacaattcat tcaagccgac accgcttcgc ggcgcggctt 8580
aattcaggag ttaaacatca tgagggaagc ggtgatcgcc gaagtatcga ctcaactatc 8640
agaggtagtt ggcgtcatcg agcgccatct cgaaccgacg ttgctggccg tacatttgta 8700
cggctccgca gtggatggcg gcctgaagcc acacagtgat attgatttgc tggttacggt 8760
gaccgtaagg cttgatgaaa caacgcggcg agctttgatc aacgaccttt tggaaacttc 8820
ggcttcccct ggagagagcg agattctccg cgctgtagaa gtcaccattg ttgtgcacga 8880
cgacatcatt ccgtggcgtt atccagctaa gcgcgaactg caatttggag aatggcagcg 8940
caatgacatt cttgcaggta tcttcgagcc agccacgatc gacattgatc tggctatctt 9000
gctgacaaaa gcaagagaac atagcgttgc cttggtaggt ccagcggcgg aggaactctt 9060
tgatccggtt cctgaacagg atctatttga ggcgctaaat gaaaccttaa cgctatggaa 9120
ctcgccgccc gactgggctg gcgatgagcg aaatgtagtg cttacgttgt cccgcatttg 9180
gtacagcgca gtaaccggca aaatcgcgcc gaaggatgtc gctgccgact gggcaatgga 9240
gcgcctgccg gcccagtatc agcccgtcat acttgaagct aggcaggctt atcttggaca 9300
agaagatcgc ttggcctcgc gcgcagatca gttggaagaa tttgttcact acgtgaaagg 9360
cgagatcacc aaggtagtcg gcaaataatg tctaacaatt cgttcaagcc gacgccgctt 9420
cgcggcgcgg cttaactcaa gcgttagaga gctggggaag actatgcgcg atctgttgaa 9480
ggtggttcta agcctcgtac ttgcgatggc atcggggcag gcacttgctg acctgccaat 9540
tgttttagtg gatgaagctc gtcttcccta tgactactcc ccatccaact acgacatttc 9600
tccaagcaac tacgacaact ccataagcaa ttacgacaat agtccatcaa attacgacaa 9660
ctctgagagc aactacgata atagttcatc caattacgac aatagtcgca acggaaatcg 9720
taggcttata tatagcgcaa atgggtctcg cactttcgcc ggctactacg tcattgccaa 9780
caatgggaca acgaacttct tttccacatc tggcaaaagg atgttctaca ccccaaaagg 9840
ggggcgcggc gtctatggcg gcaaagatgg gagcttctgc ggggcattgg tcgtcataaa 9900
tggccaattt tcgcttgccc tgacagataa cggcctgaag atcatgtatc taagcaacta 9960
gcctgctctc taataaaatg ttaggagctt ggctgccatt tttggggtga ggccgttcgc 10020
ggccgagggg cgcagccccc ggggggatgg gaggcccgcg ttagcgggcc gggagggttc 10080
gagaaggggg ggcacccccc ttcggcgtgc gcggtcacgc gccagggcgc agccctggtt 10140
aaaaacaagg tttataaata ttggtttaaa agcaggttaa aagacaggtt agcggtggcc 10200
gaaaaacggg cggaaaccct tgcaaatgct ggattttctg cctgtggaca gcccctcaaa 10260
tgtcaatagg tgcgcccctc atctgtcagc actctgcccc tcaagtgtca aggatcgcgc 10320
ccctcatctg tcagtagtcg cgcccctcaa gtgtcaatac cgcagggcac ttatccccag 10380
gcttgtccac atcatctgtg ggaaactcgc gtaaaatcag gcgttttcgc cgatttgcga 10440
ggctggccag ctccacgtcg ccggccgaaa tcgagcctgc ccctcatctg tcaacgccgc 10500
gccgggtgag tcggcccctc aagtgtcaac gtccgccccc catctgtcag tgagggccaa 10560
gttttccgcg aggtatccac aacgccggcg gccggccgcg gtgtctcgca cacggcttcg 10620
acggcgtttc tggcgcgttt gcagggccat agacggccgc cagcccagcg gcgagggcaa 10680
ccagcccggt gagcgtcgga aagggtcgac atcttgctgc gttcggatat tttcgtggag 10740
ttcccgccac agacccggat tgaaggcgag atccagcaac tcgcgccaga tcatcctgtg 10800
acggaacttt ggcgcgtgat gactggccag gacgtcggcc gaaagagcga caagcagatc 10860
acgattttcg acagcgtcgg atttgcgatc gaggattttt cggcgctgcg ctacgtccgc 10920
gaccgcgttg agggatcaag ccacagcagc ccactcgacc ttctagccga cccagacgag 10980
ccaagggatc tttttggaat gctgctccgt cgtcaggctt tccgacgttt gggtggttga 11040
acagaagtca ttatcgtacg gaatgccagc actcccgagg ggaaccctgt ggttggcatg 11100
cacatacaaa tggacgaacg gataaacctt ttcacgccct tttaaatatc cgttattcta 11160
ataaacgctc ttttctctta ggtttacccg ccaatatatc ctgtcaaaca ctgatagttt 11220
aaactgaagg cgggaaacga caatctgatc atgagcggag aattaaggga gtcacgttat 11280
gacccccgcc gatgacgcgg gacaagccgt tttacgtttg gaactgacag aaccgcaacg 11340
attgaaggag ccactcagcc ccaatacgca aaccgcctct ccccgcgcgt tggccgattc 11400
attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa 11460
ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg cttccggctc 11520
gtatgttgtg tggaattgtg agcggataac aatttcacac aggaaacagc tatgaccatg 11580
attacgccaa gctatttagg tgacactata gaatactcaa gctatgcatc caacgcgttg 11640
ggagctctcc catatcgacc tgcaggcggc cgctcgacga attaattcca atcccacaaa 11700
aatctgagct taacagcaca gttgctcctc tcagagcaga atcgggtatt caacaccctc 11760
atatcaacta ctacgttgtg tataacggtc cacatgccgg tatatacgat gactggggtt 11820
gtacaaaggc ggcaacaaac ggcgttcccg gagttgcaca caagaaattt gccactatta 11880
cagaggcaag agcagcagct gacgcgtaca caacaagtca gcaaacagac aggttgaact 11940
tcatccccaa aggagaagct caactcaagc ccaagagctt tgctaaggcc ctaacaagcc 12000
caccaaagca aaaagcccac tggctcacgc taggaaccaa aaggcccagc agtgatccag 12060

ccccaaaaga gatctccttt gccccggaga ttacaatgga cgatttcctc tatctttacg 12120
atctaggaag gaagttcgaa ggtgaaggtg acgacactat gttcaccact gataatgaga 12180
aggttagcct cttcaatttc agaaagaatg ctgacccaca gatggttaga gaggcctacg 12240
cagcaggtct catcaagacg atctacccga gtaacaatct ccaggagatc aaataccttc 12300
ccaagaaggt taaagatgca gtcaaaagat tcaggactaa ttgcatcaag aacacagaga 12360
aagacatatt tctcaagatc agaagtacta ttccagtatg gacgattcaa ggcttgcttc 12420
ataaaccaag gcaagtaata gagattggag tctctaaaaa ggtagttcct actgaatcta 12480
aggccatgca tggagtctaa gattcaaatc gaggatctaa cagaactcgc cgtgaagact 12540
ggcgaacagt tcatacagag tcttttacga ctcaatgaca agaagaaaat cttcgtcaac 12600
atggtggagc acgacactct ggtctactcc aaaaatgtca aagatacagt ctcagaagac 12660
caaagggcta ttgagacttt tcaacaaagg ataatttcgg gaaacctcct cggattccat 12720
tgcccagcta tctgtcactt catcgaaagg acagtagaaa aggaaggtgg ctcctacaaa 12780
tgccatcatt gcgataaagg aaaggctatc attcaagatc tctctgccga cagtggtccc 12840
aaagatggac ccccacccac gaggagcatc gtggaaaaag aagacgttcc aaccacgtct 12900
tcaaagcaag tggattgatg tgacatctcc actgacgtaa gggatgacgc acaatcccac 12960
tatccttcgc aagacccttc ctctatataa ggaagttcat ttcatttgga gaggacacgc 13020
tcgagacaag tttgtacaaa aaagctgaac gagaaacgta aaatgatata aatatcaata 13080
tattaaatta gattttgcat aaaaaacaga ctacataata ctgtaaaaca caacatatcc 13140
agtcactatg aatcaactac ttagatggta ttagtgacct gtagtcgacc gacagccttc 13200
caaatgttct tcgggtgatg ctgccaactt agtcgaccga cagccttcca aatgttcttc 13260
tcaaacggaa tcgtcgtatc cagcctactc gctattgtcc tcaatgccgt attaaatcat 13320
aaaaagaaat aagaaaaaga ggtgcgagcc tctttttgt gtgacaaaat aaaaacatct 13380
acctattcat atacgctagt gtcatagtcc tgaaaatcat ctgcatcaag aacaatttca 13440
caactcttat acttttctct tacaagtcgt tcggcttcat ctggattttc agcctctata 13500
cttactaaac gtgataaagt ttctgtaatt tctactgtat cgacctgcag actggctgtg 13560
tataagggag cctgacattt atattcccca gaacatcagg ttaatggcgt ttttgatgtc 13620
attttcgcgg tggctgagat cagccacttc ttccccgata acggagaccg gcacactggc 13680
catatcggtg gtcatcatgc gccagctttc atccccgata tgcaccaccg ggtaaagttc 13740
acgggagact ttatctgaca gcagacgtgc actggccagg gggatcacca tccgtcgccc 13800
gggcgtgtca ataatatcac tctgtacatc cacaaacaga cgataacggc tctctctttt 13860
ataggtgtaa accttaaact gcatttcacc agtccctgtt ctcgtcagca aaagagccgt 13920
tcatttcaat aaaccgggcg acctcagcca tcccttcctg attttccgct ttccagcgtt 13980
cggcacgcag acgacgggct tcattctgca tggttgtgct taccagaccg gagatattga 14040
catcatatat gccttgagca actgatagct gtcgctgtca actgtcactg taatacgctg 14100
cttcatagca cacctctttt tgacatactt cgggtagtgc cgatcaacgt ctcattttcg 14160
ccaaaagttg gcccagggct tcccggtatc aacagggaca ccaggattta tttattctgc 14220
gaagtgatct tccgtcacag gtatttattc ggcgcaaagt gcgtcgggtg atgctgccaa 14280
cttagtcgac tacaggtcac taataccatc taagtagttg attcatagtg actggatatg 14340
ttgtgtttta cagtattatg tagtctgttt tttatgcaaa atctaattta atatattgat 14400
atttatatca ttttacgttt ctcgttcagc tttcttgtac aaagtggtct cgaggaattc 14460
ggtaccccag cttggtaagg aaataattat tttctttttt ccttttagta taaaatagtt 14520
aagtgatgtt aattagtatg attataataa tatagttgtt ataattgtga aaaaataatt 14580
tataaatata ttgtttacat aaacaacata gtaatgtaaa aaaatatgac aagtgatgtg 14640
taagacgaag aagataaaag ttgagagtaa gtatattatt tttaatgaat ttgatcgaac 14700
atgtaagatg atatactagc attaatattt gttttaatca taatagtaat tctagctggt 14760
ttgatgaatt aaatatcaat gataaaatac tatagtaaaa ataagaataa ataaattaaa 14820
ataatatttt tttatgatta atagtttatt atataattaa atatctatac cattactaaa 14880
tattttagtt taaaagttaa taaatatttt gttagaaatt ccaatctgct tgtaatttat 14940
caataaacaa aatattaaat aacaagctaa agtaacaaat aatatcaaac taatagaaac 15000
agtaatctaa tgtaacaaaa cataatctaa tgctaatata acaaagcgca agatctatca 15060
ttttatatag tattattttc aatcaacatt cttattaatt tctaaataat acttgtagtt 15120
ttattaactt ctaaatggat tgactattaa ttaaatgaat tagtcgaaca tgaataaaca 15180
aggtaacatg atagatcatg tcattgtgtt atcattgatc ttacatttgg attgattaca 15240
gttgggaagc tgggttcgaa atcgataagc ttgcgctgca gttatcatca tcatcataga 15300
cacacgaaat aaagtaatca gattatcagt taaagctatg taatatttgc gccataacca 15360
atcaattaaa aaatagatca gtttaaagaa agatcaaagc tcaaaaaaat aaaaagagaa 15420
aagggtccta accaagaaaa tgaaggagaa aaactagaaa tttacctgca caagcttgga 15480
tcctctagac cactttgtac aagaaagctg aacgagaaac gtaaaatgat ataaatatca 15540
atatattaaa ttagattttg cataaaaaac agactacata atactgtaaa acacaacata 15600
tccagtcact atgaatcaac tacttagatg gtattagtga cctgtagtcg actaagttgg 15660
cagcatcacc cgacgcactt tgcgccgaat aaatacctgt gacggaagat cacttcgcag 15720
aataaataaa tcctggtgtc cctgttgata ccgggaagcc ctgggccaac ttttggcgaa 15780
aatgagacgt tgatcggatt tcacaactct tatacttttc tcttacaagt cgttcggctt 15840

```
catctggatt ttcagcctct atacttacta aacgtgataa agtttctgta atttctactg 15900
tatcgacctg cagactggct gtgtataagg gagcctgaca tttatattcc ccagaacatc 15960
aggttaatgg cgtttttgat gtcattttcg cggtggctga gatcagccac ttcttccccg 16020
ataacggaga ccggcacact ggccatatcg gtggtcatca tgcgccagct ttcatccccg 16080
atatgcacca ccgggtaaag ttcacgggag actttatctg acagcagacg tgcactggcc 16140
agggggatca ccatccgtcg cccgggcgtg tcaataatat cactctgtac atccacaaac 16200
agacgataac ggctctctct tttataggtg taaaccttaa actgcatttc accagtccct 16260
gttctcgtca gcaaaagagc cgttcatttc aataaaccgg gcgacctcag ccatcccttc 16320
ctgattttcc gctttccagc gttcggcacg cagacgacgg gcttcattct gcatggttgt 16380
gcttaccaga ccggagatat tgacatcata tatgccttga gcaactgata gctgtcgctg 16440
tcaactgtca ctgtaatacg ctgcttcata gcacacctct ttttgacata cttctgttct 16500
tgatgcagat gattttcagg actatgacac tagcgtatat gaataggtag atgtttttat 16560
tttgtcacac aaaaaagagg ctcgcacctc tttttcttat ttctttttat gatttaatac 16620
ggcattgagg acaatagcga gtaggctgga tacgacgatt ccgtttgaga agaacatttg 16680
gaaggctgtc ggtcgactaa gttggcagca tcacccgaag aacatttgga aggctgtcgg 16740
tcgactacag gtcactaata ccatctaagt agttgattca tagtgactgg atatgttgtg 16800
ttttacagta ttatgtagtc tgttttttat gcaaaatcta atttaatata ttgatattta 16860
tatcatttta cgtttctcgt tcagcttttt tgtacaaact tgtctagagt cctgctttaa 16920
tgagatatgc gagacgccta tgatcgcatg atatttgctt tcaattctgt tgtgcacgtt 16980
gtaaaaaacc tgagcatgtg tagctcagat ccttaccgcc ggtttcggtt cattctaatg 17040
aatatatcac ccgttactat cgtattttta tgaataatat tctccgttca atttactgat 17100
tgtaccctac tacttatatg tacaatatta aaatgaaaac aatatattgt gctgaatagg 17160
tttatagcga catctatgat agagcgccac aataacaaac aattgcgttt tattattaca 17220
aatccaattt taaaaaaagc ggcagaaccg gtcaaaccta aaagactgat tacataaatc 17280
ttattcaaat ttcaaaaggc cccaggggct agtatctacg acacaccgag cggcgaacta 17340
ataacgttca ctgaagggaa ctccggttcc ccgccggcgc gcatgggtga gattccttga 17400
agttgagtat tggccgtccg ctctaccgaa agttacgggc accattcaac ccggtccagc 17460
acggcggccg ggtaaccgac ttgctgcccc gagaattatg cagcattttt ttggtgtatg 17520
tgggccccaa atgaagtgca ggtcaaacct tgacagtgac gacaaatcgt tgggcgggtc 17580
cagggcgaat tttgcgacaa catgtcgagg ctcagcagga cctgcaggca tgcaagctag 17640
cttactagtg atgcatattc tatagtgtca cctaaatctg c                     17681
```